

ATTRIBUTION AND EVALUATION OF WORKS OF ART: THE APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES TO PAINTINGS FROM THE CIRCLE OF 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: pictorial attribution, artificial intelligence, Rembrandt

1. Attribution and authentication of works of art

The attribution of a work of art is a rather complex operation, and one can easily be misled; therefore, an in-depth study of the work in question is essential.

The difficulty also arises from the fact that, between the original work and a fake, there are almost always numerous intermediate categories, such as copies and replicas. Replicas, although executed by the same hand, present small differences that are sometimes difficult to recognize. Such works must therefore be studied at length and with great expertise: to a careful and experienced eye, quality, as the discriminating factor cannot be disregarded.

Often, in order to distinguish a replica or a copy from the original, it is necessary to perform reflectographic or radiographic analyses from which possible *pentimenti* (changes made by the artist during the execution of the work) may emerge. In this regard, without wanting to generalize, it is possible to state that, if there are no *pentimenti* in a painting, it generally means that it is not an original.

A fake, on the other hand, constitutes fraud. Forgeries have always existed: great writers such as Martial, Cicero, and Pliny testify to the fact that trade in fake art works has been a widespread phenomenon since ancient Roman times; however, at that time reproduction was nothing more than a tribute to the greatness of the artist. Only with the culture of the Enlightenment did reproductions begin to assume a negative connotation and, when, from the nineteenth century onward, works of art began to

acquire a certain economic value, authenticity became a fundamental aspect, and the improper use of signatures increased. The market for forgeries is extensive, and the corresponding economic return is considerable: according to experts, a very large percentage of works currently on the market consist of fakes.

Nevertheless, despite the paradoxes and contaminations between the two sister arts (counterfeiting and criticism), as stated by the eminent American historian Anthony Grafton, it is indisputable that “the exercise of criticism is a sign of the health and virtue of a civilization, whereas the prevalence of forgery is a sign of illness and vice”.

* Corresponding author: Lucio.colizzi@uniba.it

It therefore seems appropriate to examine the terminology used to distinguish the different degrees of certainty in the attribution of a work of art:

- **Authentic:** the work of art belongs entirely to the indicated period and to a specific author;
- **Original:** the work of art is genuinely by a specific author and presents all the stylistic characteristics of that author to be confirmed as an authentic painting with the use of diagnostic-analytical technologies;
- **Replica:** a re-edition of an original prototype executed by the author himself;
- **Copy:** a reproduction of the work executed by a different artist;
- **Attributed to:** this indication highlights that the work was produced during the lifetime of the author in question and that he is considered the most probable author;
- **Signed by:** this indication aims to guarantee the attribution to the named author, although the authenticity of the signature must still be verified;
- **School of:** the author of the work is an artist operating within the environment of the cited master or is one of his pupils;
- **Follower:** an artist who reveals certain stylistic traits attributable to a particular author or master;
- **Forgery:** the total substitution of an artifact for speculative purposes;
- **Reproduced:** a work that is identical to the original but produced through computerization.

At this point, a question arises: “What is the methodology used to evaluate the authenticity of a work of art, and how is it carried out?”

Usually, the evaluation performed by experts is subjective in nature and is based on the analysis of stylistic, aesthetic, iconographic and therefore visual elements of the work under examination. For such a subjective evaluation, a scholar’s experience is obviously essential. However, it must be supported by appropriate diagnostic technologies which, in respecting the uniqueness and irreplaceability of the cultural asset, must be non-destructive and non-invasive and possess characteristics of sensitivity, specificity repeatability and reliability.

This objective evaluation makes it possible to attribute a work, verify its authenticity and assess its state of conservation, based on the use of diagnostic methodologies, including those related to new technologies.

Ultimately, the study of a work of art must be conducted through the synergistic contribution of the diagnostician-conservator and the art historian, thus employing diagnostic and analytical technologies together with historical and philological analyses. This interdisciplinary approach constitutes the foundation for the training of the “conservator” in the field of cultural heritage enhancement within a context that brings together both scientific and humanistic expertise [1-9].

It is equally important to briefly mention the digital reproduction of a work of art [10-12]. It is not a copy – even though it may be an excellent one of a past masterpiece which has been repainted with the materials typical of the period and a technique capable of deceiving even the greatest experts – but rather, a painting that is identical to the original in size, application and material. That is because, through computerization, it is possible to obtain an artifact that is identical to the original, and whose form, colours and even “materiality” are absolute.

Since a painting is a unique and irreproducible artifact, however successful its reproduction may be, its counterfeiting must nevertheless be condemned.

At this point, several matters arise that are more ethical than aesthetic: Where has the principle of uniqueness gone? Where is its irreproducibility? And where is its inevitable market value? The simplest way to respond to these questions is that of relying

on a serious and scientific analysis of the operation. One might also consider the computerized reproduction of a work legitimate provided that its nature is clearly indicated.

The technique, however, does have practical and positive implications. Hence, it is worth noting that this sophisticated form of reproduction, for example, in the case of loans for exhibitions and subsequent returns, while preventing the “aura” of the original work from being experienced because it is not physically exhibited, nevertheless allows the original to be safely and securely preserved since it does not need to be removed from its original place of conservation.

Ultimately, if in the near future – which in reality already corresponds to the present – we have impeccable digital reproductions instead of the usual colour photographic reproductions, it can only be welcomed, because it offers the possibility of an experience that is almost identical to that of the original work. In practical terms, the continuous transfer of artworks for celebratory exhibitions can be avoided in those cases where works that cannot be safely moved are replaced by reproductions, thus completing the exhibition and avoiding not only risky transportation but also the approximate impression visitors are usually offered by ordinary photographic copies.

Given the difficulty of clearly identifying the original work, the following section presents a case study on the application of artificial intelligence (AI) techniques in the attribution of certain works of art.

2. The application of Artificial Intelligence in the field of art history: the case of the attribution of works of art from the circle of Rembrandt van Rijn

As previously mentioned, the attribution of authorship of art works has always represented one of the most complex and debated problems in art history, particularly when it concerns artists operating within workshop or school contexts characterized by strong stylistic, technical, and iconographic affinities. The case of Rembrandt van Rijn and the painters directly connected to him constitutes one of the most emblematic examples of this complexity: the presence of numerous pupils, collaborators, and imitators often makes it difficult to distinguish with certainty the hand of the master from that of his followers.

In recent years, Artificial Intelligence and machine learning techniques have begun to be explored as supporting tools for attribution analysis, thanks to their ability to identify complex visual patterns that are not immediately perceptible to the human observer. However, the application of such techniques in the field of art history raises important methodological questions, particularly in cases where the application domain presents a high degree of stylistic homogeneity and an intrinsic risk of false positives.

The present work is situated within this context and proposes an experimental analysis based on the use of the Artificial Intelligence system Luminari (owned by Atila Soares, co-author of this work) for the evaluation of the authorship of works attributed to the circle of Rembrandt. The study examines two specific cases: *Self Portrait at the Age of 34* (1640) by Rembrandt universally recognized as authentic, housed in the National Gallery, London, and used as a control reference, and *Portrait of a Man (The Auctioneer)* (ca.1658-62), a painting preserved at the Metropolitan Museum of Art in New York and the subject of an ongoing attribution debate.

Through the comparison of different machine learning and deep learning models, trained on a large dataset of works by Rembrandt and Rembrandt-related painters, the study aims to evaluate both the potential and the limitations of Artificial Intelligence in

supporting authorship attribution. The objective is not to replace historical-critical judgment, but rather to contribute to an integrated methodology in which computational analysis can critically and transparently complement humanistic expertise.

3. State of the art

The verification of authenticity and the attribution of paintings are fundamental challenges in art history and in the art market. Traditionally entrusted to the expert eye of the connoisseur, these activities require an in-depth analysis of style, materials, and provenance. In recent years, advances in AI, particularly in Deep Learning, have opened new perspectives for supporting the authentication and automatic identification of the authors of paintings [13]. From 2020 to 2025, there has been a significant increase in studies applying deep neural networks to the analysis of paintings, marking a notable growth trend in the technical and scientific literature. The following section presents a structured state of the art of the most relevant methodologies, datasets, architectures, and practical applications in this field, together with a critical synthesis of the main contributions.

Recent methodologies mainly exploit convolutional neural networks (CNNs), generative adversarial models (GANs), and, more recently, Transformers, often adapted to the specific context of paintings. Many early approaches employed CNNs pre-trained on large datasets (e.g., ImageNet) to extract features, followed by traditional classifiers. For example, Ugail et al. use a pre-trained ResNet50 as a feature extractor combined with an SVM classifier to recognize authentic paintings by Raphael [13]. This transfer learning approach has proven effective with limited data, achieving classification accuracies of around 98% in distinguishing Raphael from non-Raphael works. In general, deep CNNs (ResNet, EfficientNet, etc.) have proven capable of capturing distinctive pictorial details, such as brushstrokes, textures, and compositional structures, which help distinguish styles and authors [14].

At the same time, some studies have explored Transformer architectures for art analysis. Vision Transformers (ViT) and variants such as the Swin Transformer make it possible to capture long-range relationships within an image and to combine global context with local details [15]. A direct comparison between CNNs and Transformers on an authentication task (the case of Vincent van Gogh) showed that Transformers can achieve performance comparable to or better than traditional CNNs [15]. In one study, a pre-trained Swin Transformer model achieved more than 90% accuracy in recognizing fake Van Gogh paintings, slightly outperforming an EfficientNet-B0 model (approximately 80%) trained on the same dataset [15].

These models are also used to generate synthetic data useful for training classification models, as in the study by Ostmeyer et al., in which images generated by GANs and diffusion models improved the accuracy of classifiers in identifying forgeries [16]. Furthermore, hybrid approaches combining CNNs and GAN models have been tested to improve the reliability of the authentication process.

One of the main challenges in this field is the scarcity of well-labeled datasets. For this reason, many recent studies have developed datasets specifically dedicated to well-known artists, such as in the case of Van Gogh [16] and Rembrandt [17]. Some researchers have also exploited multispectral data, such as MA-XRF (X-ray fluorescence) datasets, combining visual information with chemical spectra to enrich the analysis and reduce false positives [18].

Artificial Intelligence has been applied in practical contexts with notable results. Boccuzzo et al. [19] developed a system capable of recognizing forgeries produced by

the well-known forger Beltracchi. Fraile-Narváez et al. [17] addressed the challenge of distinguishing between paintings by Rembrandt and those by his pupils and contemporaries, using neural networks trained on balanced datasets. These models proved effective in classifying both real and synthetic works.

Despite these advances, several technical and cultural challenges remain. The need for high-quality data, the interpretability of models, and the necessity of interdisciplinary integration are still open issues. More recent models based on Transformers and interpretable neural networks (XAI) open new perspectives for more reliable and transparent tools [15, 20]. Moreover, the use of knowledge graphs and multimodal approaches may enrich the automatic understanding of artistic contexts.

With the growth of AI-generated art, the problem of distinguishing between human and digital creations has also emerged. Velasco et al. [21] demonstrated that CNN architectures can effectively detect works created by generative systems, an issue that is likely to become increasingly relevant in the coming years.

4. Study objectives

The main objective of this study is to evaluate the ability of different AI models to distinguish between paintings belonging to the circle of Rembrandt van Rijn, with particular attention to two case studies, as mentioned earlier: a work unanimously considered authentic (Self Portrait at the Age of 34) (Figure 1) and a controversial work, *Portrait of a Man* (The Auctioneer), currently classified by the Metropolitan Museum of Art as attributed to a generic “follower of Rembrandt” (Figure 2).

Through experimentation on a dataset composed of 1,578 images from 16 Rembrandt-related artists, the study aims to verify:



Figure 1. Rembrandt, *Self Portrait at the Age of 34* (1640), oil on canvas, National Gallery, London.



Figure 2. *Portrait of a Man* (The Auctioneer), (ca. 1658-62), oil on canvas, Metropolitan Museum of Art, New York, follower of Rembrandt.

- the precision and reliability of AI models (SVM, Random Forest, CNN, Res-Net50) in multiclass classification among authors with strong stylistic similarities;
- the effectiveness of data augmentation techniques in improving predictive performance on a small-sized dataset;
- the impact of different dataset splitting strategies (training/validation/test) on model generalization;
- the possibility of attributing disputed works with greater precision, compared with a generic “school of Rembrandt” label, by proposing a more specific authorship;
- the limitations of AI in the context of intra-school attributions, where stylistic similarities exceed the differences that can be detected by supervised algorithms.

Through this analysis, the study aims to contribute to the scientific and applied debate on the use of AI in pictorial attribution, offering both methodological and experimental reflection grounded in numerical evidence and transparent comparisons.

5. Materials and methods

In the present study on the attribution of works of art, a private Artificial Intelligence program (Luminari) was used to perform tests on the authorship of two specific paintings:

- the self-portrait, unanimously recognized as an authentic work by Rembrandt, executed in 1640 and now preserved at the National Gallery in London, used in the tests as a reference for confirming authorship;
- Portrait of a Man (The Auctioneer), dated between 1658 and 1662, is still the subject of attribution controversies and preserved at the Metropolitan Museum of Art in New York. The same institution classifies the work as belonging to a generic “Follower of Rembrandt”.

During the phase of selecting the images that were to be submitted for analysis, it was decided to include Rembrandt and the most representative painters of the so-called School of Rembrandt, namely those who were trained by him directly or were deeply influenced by his style. The objective was to verify whether the models were capable of recognizing similarities or differences between authors who share an extremely rich stylistic heritage.

In the phase of preparing the images to be used in the tests, it was decided to select Rembrandt himself, naturally, together with the most likely candidates (for both cases) among those who were influenced by his style and/or imitators belonging to the “School of Rembrandt”:

1. Rembrandt van Rijn (1606-1669)	9. Aert de Gelder (1645-1727)
2. Carel Fabritius (1622-1654)	10. Willem Drost (1633-1659)
3. Gerbrand van den Eeckhout (1621-1674)	11. Reynier van Gherwen (1620-1662)
4. Govaert Flinck (1615-1660)	12. Isaac de Jouderville (1612-1646)
5. Abraham van Dijck (1635-1680)	13. Jacobus Leveck (1634-1675)
6. Samuel van Hoogstraten (1627-1678)	14. Christoph Paudiss (1630-1666)
7. Gerrit Dou (1613-1675)	15. Karel van der Pluym (1625-1672)
8. Ferdinand Bol (1616-1680)	16. Jacques des Rousseaux (1600-1638)

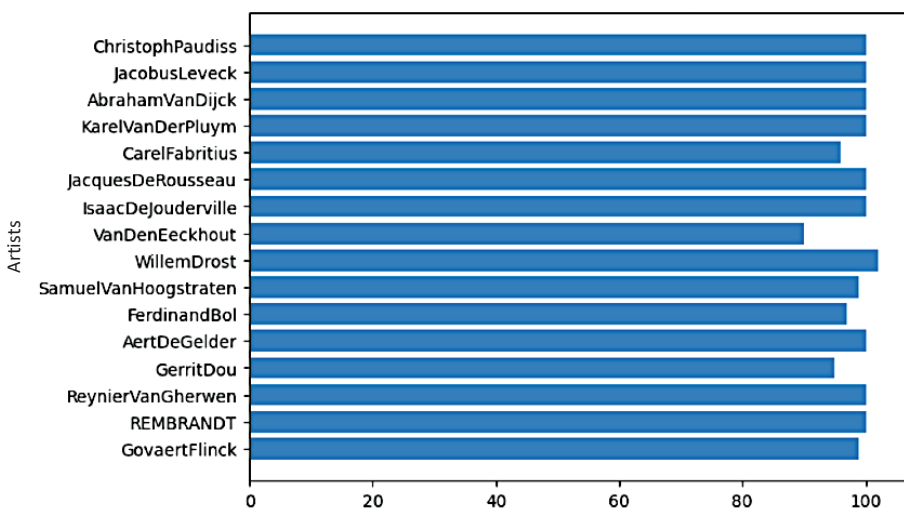
The entire experimental pipeline was implemented and tested within the Luminari platform, developed as an integrated environment for artistic attribution analysis using Artificial Intelligence. Luminari includes modules dedicated to preprocessing, training, validation, model comparison, and visualization of the results. It is based on open-source libraries (TensorFlow, Scikit-learn, PyTorch) and provides support for supervised models and deep neural networks, including pre-trained models through transfer learning.

5.1. Dataset

The dataset used in the study was constructed by selecting high-resolution digital images from publicly accessible sources, limited to the domain of portraits or half-length portraits attributable to the School of Rembrandt.

The collection consists of 1,578 images distributed across 16 authorial classes, corresponding to Rembrandt van Rijn and 16 of his historically documented followers (Table 1). Each class consists of a balanced number of images, with a range between 84 and 100 works per artist. In order to reduce bias and redundancy, digital clones, works of uncertain attribution, and images of low visual quality were excluded.

Table 1. Number of images



The original images were normalized to a size of 224×224 pixels, converted to RGB, and stored in JPEG format. An automated system for downloading and cleaning the images was implemented, followed by manual validation by two experts to ensure iconographic and stylistic consistency. Each work was labeled with the author's name according to a standardized nomenclature.

5.2. Preprocessing

Each image was resized to 224×224 pixels, converted to the RGB color space (when necessary), and normalized to the [0,1] scale by dividing by 255. The choice of this uniform format is motivated by the need for compatibility with standard convolutional neural network (CNN) architectures, particularly those pre-trained on ImageNet.

The preprocessing pipeline was automated through Python scripts (libraries: Pillow, OpenCV, NumPy) and integrated into the training sequence through the ImageDataGenerator of Keras/TensorFlow. Controlled data augmentation techniques were also implemented in order to mitigate overfitting and improve generalization:

- random rotations between -15° and $+15^\circ$
- zoom scaling of $\pm 10\%$
- horizontal flip
- brightness variation (range 0.9–1.1)
- horizontal and vertical translations up to 10%

These transformations were applied only to the training subset, while the validation and test sets were kept unchanged to ensure consistency in the evaluation.

5.3. Errors and interpretation

The in-depth analysis of the errors made by the models makes it possible to derive considerations of both a technical and an art-historical nature. In particular, it can be seen that:

- systematic confusions between Rembrandt, Bol, and Flinck may reflect a real stylistic proximity resulting from the direct influence of the master on his pupils. In this sense, the model's error may be interpreted as an indicator of workshop continuity.
- errors made in classifying works by Fabritius as Rembrandt could derive from the shared use of chiaroscuro and central compositional structures. The model appears to emphasize visual patterns more than authorial idiosyncrasies.
- the higher accuracy observed in relation to Jan Lievens may be attributed to a more independent style and to a relatively more distinctive iconographic corpus.
- the random errors observed in the CNN model compared with ResNet50 indicate a lower capacity of the former to learn deep representations: scattered confusions can be observed without a coherent pattern, with strong dependence on class balancing.

These results highlight the limitations of supervised AI when applied to contexts characterized by high stylistic density, such as the School of Rembrandt. The reliability of the system decreases as the stylistic distance between classes is reduced, suggesting that the use of AI in attribution studies should always be accompanied by art-historical expertise and by a critical interpretation of the results.


A second level of interpretation concerns the use of false positives as exploratory hypotheses: for example, if a work attributed to a follower is consistently recognized as belonging to a particular author, it may be possible to hypothesize a close relationship between the two, to be further investigated through additional technical or documentary analyses.

Error, therefore, is not only a limitation but can also become a heuristic tool, that is, a non-rigorous procedure through which a result is anticipated and that must then be formally validated within the attribution process, provided that it is properly contextualized. The following section presents the probabilities of belonging to the specific classes for both paintings examined in this scientific study.

5.4. Model results for the paintings

Table 2 reports the probabilities assigned by the classification model to the different artist classes for the two paintings analyzed: Self-Portrait (1640) and The Auctioneer. The model was trained to perform multi-class classification, in which each artist is treated as a distinct class. For each image, the system therefore produces a probability distribution over the considered classes, whose sum is equal to 1, and identifies as the predicted class the one with the highest probability.

Table 2. Model results for the specific paintings

Class (artist)	Probability Self-Portrait (1640)	Probability The Auctioneer
Class: Rembrandt	0.15	0.06
Class: SamuelVanHoogstraten	0.06	0.06
Class: KarelVanDerPluym	0.06	0.06
Class: GovaertFlinck	0.06	0.06
Class: JacquesDesRousseaux	0.06	0.06
Class: ReynierVanGherwen	0.06	0.06
Class: CarelFabritius	0.06	0.09
Class: AertDeGelder	0.06	0.09
Class: JacobusLeveck	0.06	0.06
Class: IsaacDeJouderville	0.06	0.06
Class: WillemDrost	0.06	0.06
Class: VanDenEeckhout	0.06	0.06
Class: GerritDou	0.06	0.06
Class: FerdinandBol	0.06	0.06
Class: ChristophPaudiss	0.06	0.06
Class: AbrahamVanDijck	0.06	0.06
<i>Predicted attribution of the author based on the overall probabilities</i> 	77%	71%

The table therefore allows a direct comparison of the probability distribution produced by the model in the two case studies and highlights how the system tends to place both paintings within the stylistic orbit of Rembrandt, while differentiating the class with the highest probability.

6. Discussion

The results obtained highlight strengths and some critical issues in the use of Artificial Intelligence for pictorial attribution in complex historical contexts.

On the one hand, models such as ResNet50 demonstrate that, even with non-extensive datasets, it is possible to achieve good levels of overall accuracy (up to 72%). However, this accuracy is not uniformly distributed across classes: some authors are systematically confused with each other, particularly those who share the same workshop or a clear stylistic proximity. This effect highlights a structural limitation of AI in

recognizing subtle stylistic boundaries, which are often difficult to define even for the trained human eye.

Another relevant aspect is that the models appear to learn generic visual patterns (texture, light distribution, composition) rather than deeper authorial characteristics, which may lead to correct predictions for formally incorrect reasons. This reinforces the need to combine automated tools with art-historical expertise, not as a substitute but as a support.

Moreover, the analysis of the confusion matrices suggests that false positives may represent useful clues for attributional reflection: works attributed to one author but frequently recognized as belonging to another might be re-evaluated, offering a potential exploratory tool for future studies. This perspective introduces an innovative use of classification error as a form of “predictive tension” between authors.

From a methodological point of view, the study confirms how crucial a rigorous evaluation of models is through separate datasets and appropriate metrics. The original approach, revised in light of peer review, was enriched by greater attention to cross-validation and overfitting control, thereby improving the credibility of the results.

Finally, it clearly emerges that in the absence of spectroscopic, topographic, or documentary data, AI has limited capacity to distinguish between stylistically similar authors. To overcome this limitation, for future studies, we propose the integration of multimodal approaches that combine image, spectral data, surface analysis, and contextual semantics, thus offering a more complete representation of the artwork and its context.

6.1. AI models compared

To evaluate the effectiveness of Artificial Intelligence in intra-school pictorial attribution, four models representing different families of approaches were selected:

- Support Vector Machine (SVM): a traditional non-probabilistic classifier, used with linear and RBF (Radial Basis Function) kernels. It was employed as a baseline to evaluate the discriminability of the dataset based on manually extracted HOG (Histogram of Oriented Gradients) features.
- Random Forest (RF): an ensemble of decision trees, known for its robustness on noisy and small datasets. In this case as well, the images were converted into vector features through HOG descriptors.
- Custom Convolutional Neural Network (CNN): a deep neural network composed of three convolutional blocks (conv–ReLU–maxpooling), followed by two fully connected layers. The architecture was designed to operate on images of size $224 \times 224 \times 3$ and optimized using Adam with an adaptive learning rate. The loss function used was categorical cross-entropy.
- ResNet50 with transfer learning: a model pre-trained on ImageNet, with fine-tuning of the upper layers. The classifier head was replaced with a custom two-layer multilayer perceptron (MLP), with a dropout rate of 40% to mitigate overfitting. The model was trained on a GPU (Graphics Processing Unit) for 40 epochs with early stopping based on validation loss.

All models were trained and tested on the same stratified subsets of the dataset, ensuring identical experimental conditions. The classical models (SVM and RF) were

implemented using scikit-learn, while the CNN and ResNet50 models were developed in TensorFlow/Keras.

For each model, the following indicators were recorded:

- overall accuracy (top-1)
- confusion matrix
- macro F1-score
- training time
- robustness to data augmentation

6.2. Validation strategies

To ensure a robust evaluation of the models, a validation strategy consistent with best practices in the literature on machine learning applied to medium-sized datasets was implemented.

The dataset was divided into three disjoint subsets:

- 70% Training set: used for model training.
- 15% Validation set: used for hyperparameter tuning and early stopping activation.
- 15% Test set: kept completely separate for the final evaluation of performance.

The split was performed in a stratified manner, ensuring that each author was represented proportionally in all three subsets. Randomization was performed with a fixed seed (`random_state=42`) in order to guarantee the reproducibility of the results.

During the training phase, for the deep learning models (CNN and ResNet50), an early stopping mechanism was activated by monitoring the loss on the validation set and stopping the training if no improvement was observed for 5 consecutive epochs. In addition, the model with the best weights obtained during training was retained.

Hyperparameter selection (learning rate, number of neurons, L2 regularization) was carried out through a grid search based on performance on the validation set.

For the traditional models (SVM and Random Forest), a 5-fold cross-validation was adopted on the training set, followed by final evaluation on the unseen test set. This choice made it possible to reduce the risk of overfitting and obtain a more reliable estimate of average performance.

All the metrics reported in the results refer exclusively to the test set, in order to avoid optimistic bias. The comparison between models was conducted under identical conditions of data partitioning and preprocessing.

7. Results

7.1. Metrics and visualizations

The performance of the models was evaluated using a set of standard metrics and graphical visualizations, in order to obtain a complete and comparable overview of the behavior of the algorithms.

The main metrics employed are (Table 3):

Table 3. Model / Accuracy / Precision / Recall / F1 Score / Confusion Matrix without normalization

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

- *Overall accuracy (Accuracy Top-1)*: proportion of correct predictions relative to the total number of images in the test set.
- *Macro F1-score*: harmonic mean between precision and recall calculated for each class and then averaged equally, in order to give equal weight also to less represented classes.
- *Precision and Recall per class*: reported to highlight possible imbalances between authors who are more or less easily recognized.
- *Normalized confusion matrix*: used to visualize relationships of interchange between classes, useful in the analysis of stylistic similarities perceived by the models.
- *Training time*: expressed in minutes, to evaluate the relative computational cost of the models.

The generated visualizations include:

- comparative bar charts for model accuracy and F1-score.
- heatmaps of the confusion matrix for each model.
- learning curves (accuracy and loss on train/validation sets) for the deep learning models.

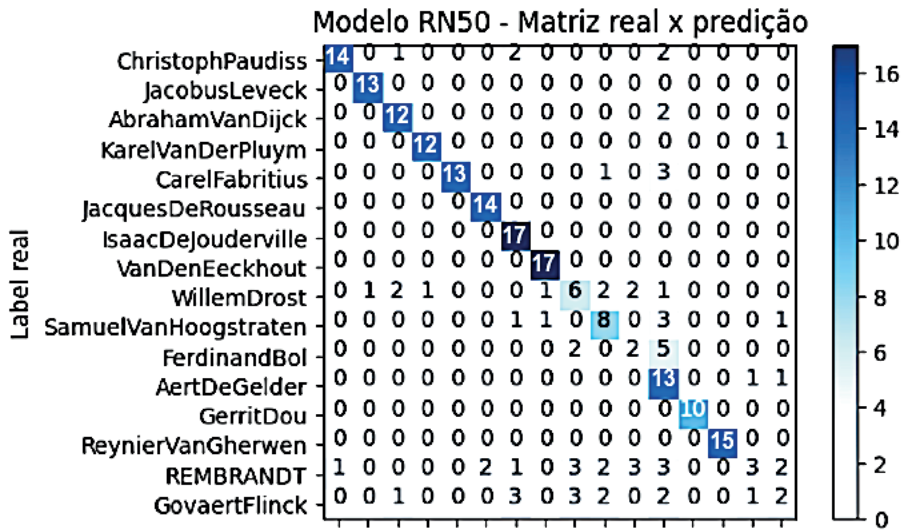
The visualizations were produced using Matplotlib and Seaborn. To ensure replicability, all analysis scripts were documented and included as supplementary material.

7.2. Analysis of the confusion matrix

The confusion matrices obtained for the tested models reveal important information about the behavior of the networks during the classification phase, particularly with regard to confusions between authors with similar pictorial styles.

For the ResNet50 model (Table 4), which achieved the highest overall performance (top-1 accuracy = 72%), the following recurring patterns can be observed:

Table 4. RN50 model – matrix reality × prediction / true label



- High confusion between Ferdinand Bol and Govert Flinck, with more than 20% of the works of the latter classified as belonging to the former. This suggests a strong stylistic overlap perceived by the model.
- Some works by Carel Fabritius are incorrectly classified as Rembrandt, possibly due to similarities in composition and the dramatic use of light.
- The model distinguishes more effectively painters who are stylistically more independent, such as Jan Lievens and Nicolaes Maes, which show precision values above 80%.

In the case of the custom CNN model, the confusion matrix appears less structured and noisier. The confusions are more widely distributed, with a reduction in average precision per class and a tendency to classify many images toward the majority classes, indicating overfitting on more represented authors.

The SVM and Random Forest models, although less performant in absolute terms, showed some consistency with the confusions observed in the deep models, especially in the Rembrandt–Bol–Flinck cluster.

The analysis of such errors is crucial in attribution studies, where the stylistic distance perceived by the model may offer alternative hypotheses or indicate relationships between workshops.

In the following section 8, a reflection is presented on how these results can be interpreted from an art-historical perspective and to what extent the model’s errors may suggest proximity or divergence among the artists considered.

8. Conclusion

The present study comparatively analyzed the effectiveness of different machine learning models for the task of pictorial attribution within the circle of Rembrandt, highlighting both advantages and some critical issues in the use of AI in contexts characterized by high stylistic density. The results indicate that, although models such as ResNet50 achieve good average levels of accuracy, performance can vary significantly depending on the stylistic proximity between authors, making automatic attribution particularly difficult in intra-school contexts.

The study also confirmed the importance of rigorous evaluation through cross-validation techniques, controlled dataset splits, and robust metrics such as macro F1-score and the confusion matrix. The methodological contribution includes a replicable pipeline, improved on the basis of two peer reviews, which can be reused or extended in other similar attribution cases. On the other hand, it is equally true that other aspects could contribute to the completeness of the research: spectroscopic data, model interpretability, the uniformity of the dataset with respect to the actual museum corpus, and the possibility of distinguishing declared forgeries from known originals. These limitations provide a clear horizon for improvement and open the way to future research directions.

Our work suggests that Artificial Intelligence, if properly guided, can play a relevant role in pictorial attribution, but it must be used within a solid, interdisciplinary, and transparent methodological framework. The most promising applications appear to lie not in replacing human expertise, but in enhancing its analytical capabilities, suggesting hypotheses, highlighting unexpected similarities, or strengthening uncertain attributions with qualitative, quantitative, and replicable arguments.

References

- [1] Lorusso, S., Vandini, M., Matteucci, C., Tumidei, S., & Campanella, L. (2003). Anamnesi storica ed indagine diagnostica del dipinto ad olio su tavola "Madonna con bambino e Santi Girolamo e Caterina da Siena" attribuibile a Domenico Beccafumi (1486-1551). *Quaderni di Scienza della Conservazione*, 47-68.
- [2] Lorusso, S., Matteucci, C., Natali, A., & Visco, V. (2006). Indagini diagnostiche e valutazione storico-stilistica del dipinto ad olio su tela raffigurante Uomo che legge a lume di lucerna. *Quaderni di Scienza della Conservazione*, 6, 117-137.
- [3] Lorusso, S., Matteucci, C., & Natali, A. (2006). Sulla attribuzione e sullo stato di conservazione di tre dipinti: anamnesi, confronto, valutazione. In *Lo stato dell'arte* (pp. 679-688). Nardini.
- [4] Lorusso, S., Natali, A., Matteucci, C., Bertolino, R., & Tumidei, S. (2007). Diagnostic investigation and historical-stylistic evaluation of oil painting on metal board. example. in *Conservation Science in Cultural Heritage*, Bologna, Pitagora Editrice, 7/2007.
- [5] Lorusso, S., Matteucci, C., Natali, A., Chinni, T., & Solla, L. (2011). The Shroud of Turin between history and science: an ongoing debate. *Conservation Science in Cultural Heritage*, 11, 113-152.
- [6] Lorusso, S., Matteucci, C., Natali, A., & Apicella, S. A. (2013). Traditional and non-traditional, innovative and ephemeral materials and techniques in today's cultural heritage. *Russian Chemical Bulletin*, 62(7), 1671-1681.

- [7] Lorusso, S., Natali, A. «Le diverse tipologie di riproduzione nell'arte in "Il diritto. La Protezione del patrimonio artistico",» Ginevra-Milano, a cura di Negri-Clementi G., Stabile S., Skira, ISBN 978 -88-572-2627-9, pp. 155-167.
- [8] Lorusso, S., & Natali, A. (2015). Mona Lisa: a comparative evaluation of the different versions and their copies. *Conservation Science in Cultural Heritage*, 15, 57-109.
- [9] Lorusso, S., Natali, A., Isbouts, J. P., Asmus, J., & Parfenov, V. (2021). A second Mona Lisa?: challenges of attribution and authentication and various possibilities for evaluating a work of art.
- [10] Lorusso, S., & Mantovani, M. (2023). Synergy, technology and identity in art, facing the challenge of artificial intelligence. *Conservation Science in Cultural Heritage*, 23, 21-38. Gangemi Editore, Roma,
- [11] Lorusso, S., Colizzi, L., & Adamo, T. (2024). Artificial intelligence and digital reproduction in art. *Conservation Science in Cultural Heritage*, 23, 42-51.
- [12] Lorusso, S., Gugliermetti, L., & Cinquepalmi, F. (2025). Digital art as a future vehicle of contemporaneity: NFTs (Non-Fungible Tokens). *Conservation Science in Cultural Heritage*, 1(24), 393.
- [13] Ugail, H., Stork, D. G., Edwards, H., Seward, S. C., & Brooke, C. (2023). Deep transfer learning for visual analysis and attribution of paintings by Raphael. *Heritage Science*, 11(1), 1-15.
- [14] Li, W. (2025). Enhanced automated art curation using supervised modified CNN for art style classification. *Scientific Reports*, 15(1), 7319.
- [15] Schaerf, L., Postma, E., & Popovici, C. (2024). Art authentication with vision transformers. *Neural Computing and Applications*, 36(20), 11849-11858.
- [16] Ostmeyer, J., Schaerf, L., Buividovich, P., Charles, T., Postma, E., & Popovici, C. (2024). Synthetic images aid the recognition of human-made art forgeries. *Plos one*, 19(2), e0295967.
- [17] Narváez, M. F., Olivenza, I. S., & McGowan, N. (2022). Painting Authorship and Forgery Detection Challenges with AI Image Generation Algorithms: Rembrandt and 17th Century Dutch Painters as a Case Study. *International Journal of Interactive Multimedia and Artificial Intelligence*, 7(7), 7-13.
- [18] Preisler, Z., Andolina, R., Busacca, A., Caliri, C., Miliani, C., & Romano, F. P. (2024). Deep learning for enhanced spectral analysis of MA-XRF datasets of paintings. *Science Advances*, 10(39), eadp6234.
- [19] Boccuzzo, S., Meyer, D. D., & Schaerf, L. (2024, September). Art forgery detection using Kolmogorov arnold and convolutional neural networks. In *European Conference on Computer Vision* (pp. 187-201). Cham: Springer Nature Switzerland.
- [20] Zanardelli, M., Guerrini, F., Leonardi, R., & Adami, N. (2023). Image forgery detection: a survey of recent deep-learning approaches. *Multimedia Tools and Applications*, 82(12), 17521-17566.
- [21] De Goma, J. C., Olivar, J. B. L., Perez, H. L., Velasco, S. C. E., & Villaluz, A. C. (2024, September). Art Authentication: A Comparative Analysis of Convolutional Neural Network (CNN) Architectures for Detecting AI-Generated and Human-made Digital Artworks. In *Proceeding of the 2024 5th Asia Service Sciences and Software Engineering Conference* (pp. 147-152).

Biographical Notes

Lucio Colizzi is a researcher at the University of Bari Aldo Moro (Italy), working in the field of computer science and software engineering. His research activity focuses on distributed and cooperative systems, Internet of Things architectures, cyber-physical systems, digital twins, and decision support systems, with applications in agritech, environmental monitoring, and cultural heritage.

He has been the scientific coordinator of several important research projects in the field of technologies for Cultural Heritage, including SIDART (Sistema Integrato per la Diagnostica dei Beni Artistici), devoted to the development of integrated diagnostic systems for the analysis and conservation of artworks; BLU-ARCHEOSYS, focused on advanced technological solutions for the investigation and monitoring of underwater archaeological heritage; and IT@CHA (Tecnologie Italiane per applicazioni avanzate nei Beni Culturali), aimed at the development and experimentation of innovative digital technologies for the study, protection, management, and valorization of cultural heritage.

His research integrates sensing technologies, advanced software architectures, artificial intelligence, and digital modelling methodologies for the monitoring, analysis, and sustainable management of complex cultural and environmental systems. He has published several scientific papers on the application of Information and Communication Technologies (ICT) to the study, diagnostic analysis, monitoring, and digital enhancement of cultural heritage assets.

Átila Soares da Costa Filho has a degree in Industrial Design from the Pontifical Catholic University of Rio de Janeiro (Brazil) with graduate degrees in History, Art History, Philosophy, Anthropology, Sociology, Archaeology and Heritage. The author of five books and numerous articles published in more than 100 countries, he is a member of the Scientific Committee of the Mona Lisa Foundation (Zurich), the Leonardo da Vinci Foundation (Milan) and the National Committee for the Enhancement of Historical, Cultural and Environmental Heritage (Rome). (Cf. <https://professoratila-soares.weebly.com/>)

Salvatore Lorusso formerly a full Professor at the University of Bologna, is currently a Foreign Member of the Russian Academy of Natural Sciences, Visiting Professor of the Faculty of Arts, Lomonosov Moscow State University, Russia and Emeritus Professor of the Cultural Heritage Institute of Zhejiang University (China) and Direttore Generale of the Accademia della Cultura Enogastronomica; formerly a Visiting Professor at the Academy of Social Science of Zhejiang University, China; formerly vice-president and Councillor of the Società Italiana per il Progresso delle Scienze (SIPS, established in 1839). His biography appears in the 2016 Marquis Edition of Who's Who in the World. He is the author of over 430 publications in national and international journals and 25 volumes and monographs covering commodity science, cultural heritage and environmental issues. In 1997, he founded the Diagnostic Laboratory for Cultural Heritage at the Ravenna Campus of the University of Bologna and remained head of the Laboratory for eighteen years. In 2001 he founded the historical-technical Journal "Conservation Science in Cultural Heritage" of which he is Editor-in-Chief. His scientific work deals mainly with the study of the "system: artifact-environment-biota" and diagnostic, analytical, technical and economic evaluation within the context of the protection and valorization of cultural and environmental heritage.

Summary

The present study analyzes the application of Artificial Intelligence techniques to the attribution of authorship of paintings belonging to the circle of Rembrandt van Rijn, a context characterized by strong stylistic similarities. Using the proprietary system Luminary, two case studies were examined: a self-portrait by Rembrandt universally recognized as authentic and Portrait of a Man (The Auctioneer), a work of controversial attribution.

The analysis is based on a dataset of 1,578 images referring to 16 painters of the School of Rembrandt. Different machine learning and deep learning models were compared, including Random Forest, Support Vector Machine, and convolutional neural networks, with particular attention to a ResNet50 pre-trained on ImageNet. The latter showed the best performance, with an overall accuracy of approximately 72%, a significant value in a domain characterized by high authorial ambiguity.

The results highlight both the potential of AI in supporting attribution analysis and the intrinsic limitations in cases of strong stylistic proximity. The introduction of a corrective index related to the School of Rembrandt allows a more cautious and informed evaluation, confirming the authenticity of the self-portrait and suggesting, for The Auctioneer, a non-conclusive attribution within the circle of Aert de Gelder.