

AI-based technologies in the authentication of fine art: toward a hybrid epistemology of cultural trust

Atanasova D. Y.

Sofia, Bulgaria

E-mail: yankovad@gmail.com

Abstract: *The authentication of fine art has customarily relied on expert connoisseurship, material analysis, and provenance research. In recent years, artificial intelligence (AI) and AI-based technologies have appeared as significant tools in this domain, enabling new forms of algorithmic evidence, probabilistic reasoning, and large-scale pattern recognition. This paper examines how AI-based systems support museums, galleries, collectors, and private institutions in authenticating fine art paintings. It argues that AI does not replace human expertise but establishes a hybrid epistemic framework in which algorithmic forensics and art-historical knowledge co-produce authenticity. The study analyses key technological approaches, institutional applications, epistemological implications, and structural drawbacks, positioning AI as a catalytic agent in reconfiguring trust, authority, and knowledge production within the contemporary art ecosystem.*

KEYWORDS: ARTIFICIAL INTELLIGENCE, ART AUTHENTICATION, HYBRID EPISTEMOLOGY, PROVENANCE ANALYSIS, CULTURAL ANALYTICS

1. Introduction

Authentication constitutes the epistemic and economic foundation of the fine art ecosystem, determining not only cultural value but also legal status, market price, and institutional legitimacy [10]. Traditional authentication methods—connoisseurship, material analysis, and provenance research—now face growing challenges from increasingly sophisticated forgeries and the global expansion of art markets, which exceed human analytical capacity.

Recent advances in artificial intelligence have introduced computational paradigms capable of analysing visual, material, and historical characteristics at unprecedented scale and precision [13]. However, the assimilation of algorithmic systems into cultural heritage raises fundamental epistemological questions about the nature of authenticity itself.

This paper contends that AI-based technologies not just supplement but fundamentally reshape the process of authenticating fine art. We advance a “hybrid epistemology of cultural trust”, which systematically integrates computational forensics, expert art-historical analysis, and institutional authority to establish authenticity. Rather than relying on a single methodological approach, this model coordinates algorithmic and human assessment across technical, scholarly, and organisational domains. While it enhances rigour and transparency, it also introduces possible challenges, including diminished expert authority, excessive dependence on opaque probabilistic methods, and reinforcement of organisational biases. Therefore, this system highlights the importance of continuous critical assessment of the limitations of merging AI with established authentication practices.

To operationalise this hybrid model, museums and universities can implement concrete measures at multiple levels. Staff training programs should be updated to include basic AI literacy, ensuring that curators, conservators, and registrars are familiar with the principles and limitations of machine learning tools. Institutions can revise authentication workflows to incorporate AI analyses as an evidentiary layer that complements traditional review panels, encouraging multidisciplinary collaboration. Policy updates might require transparent documentation of AI-supported decisions and regular audits of algorithmic outputs to promote accountability. Furthermore, ethical guidelines should be introduced to guard against bias, mandate data diversity, and clarify the interpretive role of human experts alongside automated assessments. By systematically embedding these changes, institutions will be better equipped to integrate AI into everyday authentication practice while upholding scholarly rigour and cultural trust.

2. Results and Discussions

2.1. Technological Foundations and Methodologies

- *Computer Vision and Deep Learning.* Modern AI authentication systems employ convolutional neural networks (CNNs) and visual transformers to analyse micro-level visual features, including brushstroke dynamics, colour distributions, compositional geometry, and textual patterns [5]. Recent studies show that hybrid CNN-transformer models attain improved performance by extracting both local brushstroke features and global compositional structures, with classification accuracies frequently ranging from 80% to 95% on controlled datasets [13]

The breakthrough lies in converting subjective stylistic judgement into assessable computational data, producing probabilistic outputs such as “92% probability of Rembrandt authorship” rather than binary determinations [8].

- *Multispectral and Hyperspectral Analysis.* AI-enhanced spectral imaging enables non-invasive examination through X-ray fluorescence (XRF), infrared reflectography (IRR), and hyperspectral imaging. Machine learning methods can detect anachronistic pigments, modern binders in historical works, hidden underdrawings, and restoration layers with reliability exceeding 85% under regulated conditions [3].

ASML's Impasto Project illustrates this approach, creating 3D digital twins of Vincent van Gogh's artworks with 100 gigabytes of data per scan, utilising nanometre-capable chipmaking technology to generate "Google Maps for paintings" [12].

- *Provenance and Network Analysis.* Graph-based AI systems construct knowledge graphs from auction records, exhibition catalogues, museum databases, and archival photographs. These systems can automatically detect missing provenance periods, suspicious ownership patterns, and stylistic clusters associated with known forgers, thereby raising anomaly-detection efficiency by approximately 30-40% across big datasets [9].

2.2. Institutional Applications and Case Studies

- *Museums.* Major museums use AI for broad collection analysis, but implementation reveals varying institutional priorities that shape the scope and impact of technological adoption. The Rijksmuseum applies residual neural networks to identify large-scale patterns across entire collections, emphasising a macro-level understanding of organisational trends and misattributions. In contrast, the Metropolitan Museum of Art incorporates AI more narrowly into conservation workflows for individual objects,

focusing on micro-level technical analysis and object-specific decision-making [6]. The distinction between the Rijksmuseum's collection-level approach and the Metropolitan Museum's object-level focus highlights how institutional aims shape both the strategic aims and the epistemic contributions of AI integration. Despite this contrast, both institutions face comparable challenges: ensuring unbiased training data, preventing cataloguing errors, and managing the interpretability of algorithm-generated outputs. These cases underscore the significance of aligning innovative AI implementation with established scholarly practices and shared curatorial authority to promote institutional effectiveness and credibility.

- *Auction Houses and Galleries.* Sotheby's and Christie's have incorporated AI-driven technical imaging and machine learning into their due diligence procedures, primarily aiming to enhance pre-acquisition risk assessment and market intelligence. While these integrations enable auction houses to detect potential forgeries with increased accuracy and efficiency, they also introduce new complexities into the authentication process. For instance, the reliance on algorithmic outputs may shift the locus of expertise from traditional specialists to technology providers, potentially redefining what is considered authoritative authentication. Moreover, there is a risk that overreliance on AI-generated evidence could obscure subtle interpretive or contextual nuances that human experts might discern. Despite these concerns, auction houses have leveraged these technologies to reduce reputational risk and bolster buyer confidence, directly benefiting stakeholders such as collectors, consignors, and institutional clients by providing greater assurance in the authenticity and market value of works. However, these developments also necessitate ongoing scrutiny regarding how algorithmic opacity, data representativeness, and potential biases might impact the broader art market. Thus, although AI-based implementations have the potential to increase trust and inform more rational economic decision-making in speculative markets, they simultaneously require critical evaluation to ensure that enhanced efficiency does not compromise the epistemic integrity of authentication practices [2].

Table 1: Comparative Performance of Authentication Approaches

Connoisseurship	70-85%	Contextual judgment	Subjectivity
Material Analysis	80-95%	Scientific rigour	Expensive, slow
AI Image Analysis	80-92%	Pattern detection at scale	Data dependence
Hybrid Approach	90-97%	Complementary validation	Integration complexity

- *Private Collections and Commercial Platforms.* Commercial platforms such as Art Recognition illustrate the market demand for scalable computational attribution. In 2023, Art Recognition was commissioned by a private Swiss collector to assess the authenticity of a portrait attributed to Amedeo Modigliani. Utilising their proprietary AI system, trained on verified Modigliani works, the platform analysed high-definition images and identified differences in brushwork and design features, resulting in a revised attribution [4]. The AI's rapid visual analysis detected subtle variations in brushstroke rhythm and pigment layering that typically require extensive microscopic examination by human experts. The probability analysis assigned a significantly lower likelihood of

Modigliani authorship, casting doubt on the painting's authenticity and prompting further scholarly debate regarding the evidentiary weight of AI results compared to provenance documentation and stylistic evaluation. A critical synthesis reveals that AI provided a novel form of evidence—quantifiable and replicable—that augmented but did not replace traditional authentication methods. Rather than serving as a final arbiter, the AI's findings highlight the probabilistic and provisional nature of computational analysis, emphasising the need for art historians and domain experts to scrutinise both the methodology and implications of algorithmic judgments. This case exposes the epistemological complexity of integrating AI with human expertise. While AI enabled efficient triage and revealed previously undetected features, its output depended on the completeness of reference datasets and required contextual validation within disciplinary frameworks. Ultimately, the Modigliani attribution demonstrates that AI-generated reports should function not as unilateral determinants but as catalysts for multidisciplinary dialogue, reinforcing the necessity of critical engagement and iterative review within evolving art authentication practices.

To address the broader international scope of art authentication, some platforms have begun integrating non-Western and underrepresented art traditions into their datasets, broadening institutional relevance beyond Western-centric collections. The recent pilot project involving Ming dynasty porcelain highlights this shift: AI algorithms, trained on digitised records and multispectral imagery sourced from Chinese museums and private collections, were used to authenticate ceramic works. The project's successful differentiation between authentic Ming pieces and later reproductions—based on subtle kiln signatures and glaze compositions unique to specific regions—demonstrates that AI methods can be tailored to the material and stylistic complexities of globally diverse artefacts, not solely Western paintings. Nevertheless, persistent limitations, such as the scarcity of high-quality reference images and the underrepresentation of localised styles, challenge the inclusivity and generalizability of these systems, revealing a fundamental constraint of the broader hybrid epistemological model. The Ming porcelain pilot, therefore, not only illustrates the expanded potential for AI-supported authentication in non-Western contexts but also exemplifies how the diversity and representativeness of underlying datasets ultimately limit the effectiveness of hybrid approaches. This case demonstrates that without deliberate efforts to ensure inclusivity at the data-collection and methodological levels, the promise of a truly global hybrid framework remains constrained, reinforcing the need for institutions to address structural data biases when integrating algorithmic evidence with human expertise on a worldwide scale.

Table 2: Institutional AI Applications in Art Authentication

Rijksmuseum	Computer vision	Attribution research	Collection-wide analysis
Sotheby's	Technical imaging + ML	Due diligence	Risk reduction
Christie's	Data analytics	Market intelligence	Pricing confidence
Getty Research Institute	Provenance data transformation	21st-century research	Enhanced accessibility
Art Recognition	CNN authentication	Collector support	Pre-purchase screening

2.3. The Hybrid Epistemological Framework

- *From Authority to Statistical Coherence.* Artificial intelligence shifts the concept of authenticity from reliance on expert authority to recognition of statistical patterns. Rather than asserting absolute truth, AI generates probabilistic knowledge derived from empirical data, thereby framing authenticity as a system-based process rather than solely expert opinion.

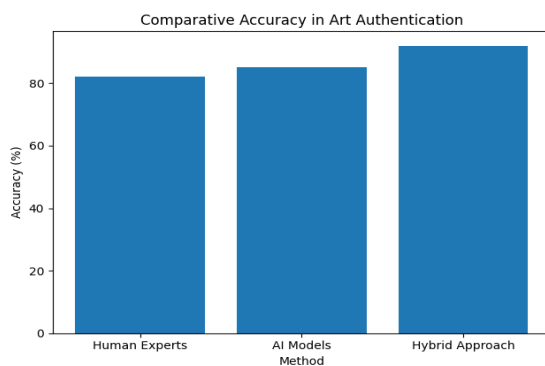
- *Augmented Connoisseurship Model.* "Augmented connoisseurship" refers to a hybrid epistemic system in which:

- Within this framework, AI provides computational proof by generating quantitative evidence through the analysis of large datasets and the application of algorithmic models.

- Experts provide contextual interpretation.
- Institutions provide legitimacy validation.

Authenticity emerges from interactions among datasets, algorithms, and institutional authentication procedures rather than from any single authority [10].

Fig. 1 (Hybrid Epistemology Diagram)



- *Cultural Trust Networks.* The Getty Research Institute's provenance data initiative embodies the hybrid model by directly combining computational innovation with established scholarly protocols. By digitising and consolidating provenance records into an accessible database, the project not only increases data transparency but also enables cross-institutional research without jeopardising curatorial oversight or scholarly review standards [7]. This concentrated approach illustrates how digital technology can expand research capacity and accessibility while retaining the strict evaluation required for data dependability and interpretive validity. In parallel, the Centre for Art Law's Framework for Responsible Use of AI in Art Authentication provides structured guidelines that integrate practical, legal, and ethical considerations, ensuring that new AI tools are introduced without undermining institutional confidence and stewardship in authentication processes [14].

2.4. Constraints and Difficulties

- *Structural Boundaries.*

- Despite impressive technical progress, AI confronts major limitations:

- **Dataset Bias:** Systems trained on historically attributed works may reproduce and amplify existing misattributions
- **The Perfect Forgery Problem:** Skilled forgeries may statistically outperform genuine experimental works which diverge from established stylistic norms
- **Ontological Limits:** AI cannot access artistic intention, symbolic meaning, or cultural context

- *Ethical and Methodological Concerns.* These challenges emphasise the vital need for a concise set of moral guidelines for the use of AI in art authentication. Primary recommendations include ensuring data representativeness to avoid marginalising non-Western artistic traditions, increasing transparency in algorithmic processes to address the opacity of deep learning, and implementing routine multidisciplinary oversight.

To implement these recommendations, institutions can take several concrete actions. First, museums and galleries should establish dedicated ethics committees to oversee the integration of AI tools, review algorithms for fairness, and evaluate their impacts on established authentication procedures. Additionally, acquisition protocols and cataloguing standards should be updated to require that provenance documentation, artists' records, and technical analyses include data fields relevant to AI-assisted evaluation and are maintained in formats compatible with computational analysis.

Institutions ought to establish clear ethical codes, apply mechanisms to correct past biases inside datasets, and avoid treating probabilistic AI outputs as definitive judgments. Routine training programs can be introduced to increase AI literacy among curators, conservators, and administrative staff, ensuring informed interpretation of AI-generated results. Regular audits and transparent reporting of AI-supported decisions further promote accountability. These specific measures, when adopted systematically, jointly aim to promote fairness, accountability, and equity, safeguarding the ethical duties of stakeholders across the global art ecosystem.

2.5. Upcoming Directions and Implications

The future trajectory of AI in art authentication is expected to involve multimodal models that analyse high-resolution visual imagery, material composition data from multispectral scanning, and digitised archival or textual documentation in combination. Implementing these innovations will require streamlined protocols for data capture, standardised metadata annotation, and systematic cross-linking of diverse sources. Advances in explainable AI, including saliency attribution maps and transparent rule-based decision trees, will enhance methodological transparency by making algorithmic reasoning accessible for review by both art historians and non-specialist stakeholders. This transparency is crucial for interpretive reliability, particularly as collaborative attribution platforms incorporate expert validation and iterative feedback. As real-time consensus-building and error-correction become more prominent, new demands will arise for effective interoperability between human and artificial agents, shifting the methodological emphasis from static individual judgment to ongoing, multidisciplinary committee-based review.

These recommendations target both institutional and policy design levels within the art authentication ecosystem. Institutionally, museums, galleries, and private collections should revise acquisition protocols and record-keeping practices to ensure compatibility with AI systems, adopt standardised metadata and imaging formats, and provide staff training in data management and AI analysis. At the policy level, regulatory bodies should develop guidelines promoting responsible, unbiased data use and equitable access to AI resources. Sector-wide initiatives, such as establishing open, multilingual metadata repositories and interoperable curation standards, will enhance participation, especially for underrepresented collections [11]. Additionally, harmonising data-sharing and funding practices across national borders will reduce disparities in access and representation. Collectively, these measures address practical stakeholder needs and core technical and ethical priorities essential for implementing a hybrid epistemological model for AI-powered art authentication [1].

Blockchain-supported certification systems, while not AI themselves, complement computational authentication by providing tamper-resistant provenance records and digital certificates anchored to distributed ledgers, boosting disclosure and long-term traceability.

3. Conclusion

To summarise, this paper makes an original contribution by articulating and substantiating a hybrid epistemological model for art authentication that uniquely unites computational forensics, expert interpretation, and institutional validation. The research advances the field by not only synthesising the roles of AI and human expertise, but also by proposing a structured partnership that clarifies how each component contributes functionally to authentication processes. Via systematic analysis of the limitations and synergies inherent in uniting these elements, the paper argues that authenticity is most reliably established through this combined framework rather than through automation alone. This approach offers a novel foundation for both scholarly discussion and institutional policy by enhancing technical accuracy, protecting the integrity of cultural heritage, and strengthening organisational accountability and societal trust. Thus, the proposed synthesis establishes a strong and flexible blueprint for addressing emerging challenges in contemporary art authentication.

The proposed hybrid epistemological model, which integrates computational forensics, expert interpretation, and institutional validation, delivers a sustainable framework for cultural heritage institutions navigating technological transformation while continuing scholarly rigour and cultural trust. Success depends not on automation but on the careful orchestration of human-computer collaboration that preserves the interpretive richness of art-historical scholarship while leveraging AI's analytical strengths. To accelerate the adoption and effectiveness of this hybrid model, policy initiatives and funding opportunities should be prioritised. Policymakers and funding agencies can support institutions by providing resources to upgrade digital infrastructure, develop standardised protocols for AI integration, and facilitate cross-institutional collaboration. Dedicated funding streams for research, staff training, and data curation are essential to ensure equitable access and to address disparities between well-resourced and underrepresented collections. Sector-wide policy frameworks that encourage ethical AI practices and ongoing evaluation will further strengthen the transition to a transparent and accountable authentication ecosystem.

The future of art authentication will be defined by institutions that master this integration, creating transparent, accountable systems in which artificial intelligence functions as a strong evidentiary tool inside broader ecosystems of cultural knowledge and trust.

4. References

1. A Adorni, G. & Bellini, E. (2025). Towards a Manifesto for Cyber Humanities: Paradigms, Ethics, and Prospects. arXiv preprint. <https://doi.org/10.48550/arXiv.2508.02760>, Accessed: 10.02.2026
2. Lawson-Tancred, J., (2025). AI and the Art Market (Hot Topics in the Art World). <https://oceanofpdf.com/authors/jo-lawson-tancred/pdf-epub-ai-and-the-art-market-hot-topics-in-the-art-world-download/>, Accessed: 12.02.2026
3. Polak, A., Kelman, T., Murray, P., Marshall, S., (2017). Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. https://www.researchgate.net/publication/314670114_Hyperspectral_imaging_combined_with_data_classification_techniques_as_an_aid_for_artwork_authentication, Accessed: 15.02.2026
4. Dobbs, T., Nayeem, A., Cho, I. & Ras, Z. (2023). Contemporary Art Authentication with Large-Scale Classification. *Big Data Cogn. Comput.* <https://doi.org/10.3390/bdcc7040162>, Accessed 01.02.2026
5. Elgammal, A., Mazzone, M., Liu, B., Kim, D., & Elhoseiny, M. (2018). The Shape of Art History in the Eyes of the Machine. <https://www.researchgate.net/publication/322694894>, Accessed: 01.02.2026
6. Expert Systems with Applications. (2022). On art authentication and the Rijksmuseum challenge: A residual neural network approach, 202. <https://doi.org/10.1016/j.eswa.2022.116933>, Accessed: 05.02.2026
7. Getty Research Institute. (May 1, 2025). Getty Transforms Art Provenance Data to Support 21st Century Research. Getty News. <https://www.getty.edu/news/getty-transforms-art-provenance-data-to-support-21st-century-research/>, Accessed: 07.02.2026
8. Johnson, C. R., Hendriks, E., Berezhnoy, I., et al. (2008). Image processing for artist identification. <https://ericpostma.nl/publications/Johnsonetal2008.pdf>, Accessed: 11.02.2026
9. Münster, S., et al. (2019). Digital cultural heritage meets digital humanities. https://www.researchgate.net/publication/335417192_DIGITAL_CULTURAL_HERITAGE_MEETS_DIGITAL_HUMANITIES, Accessed: 07.02.2026
10. Silva, R. S., Lotfi, A., Ihianle, I. K., Shahtahmassebi, G., & Bird, J. J. (2024). ArtBrain: An Explainable End-to-End Toolkit for Classification and Attribution of AI-Generated Art and Style. arXiv preprint. <https://doi.org/10.48550/arXiv.2412.01512>, Accessed: 10.02.2026
11. Thiel, S. & Bernhardt, J. C. (2024). AI in Museums: Reflections, Perspectives and Applications. transcript publishing. https://library.oapen.org/bitstream/handle/20.500.12657/87430/1/9783839467107.pdf?trk=public_post_comment-text, Accessed: 02.02.2026
12. Tom's Hardware. (June 14, 2025). ASML's Impasto Project creates 3D digital twins of Vincent van Gogh's art with 100 gigabytes of data per scan. <https://www.tomshardware.com/3d-printing/3d-scanning/asmls-impasto-project-creates-3d-digital-twins-of-vincent-van-goghs-art-with-100-gigabytes-of-data-per-scan-nanometer-capable-chipmaking-tech-used-to-create-google-maps-for-paintings>
13. Wang, Z. & Song, H. (2025). A Fusion Model for Artwork Identification Based on Convolutional Neural Networks and Transformers.. <https://doi.org/10.48550/arXiv.2502.18083>, Accessed: 08.02.2026
14. Centre for Art Law. (2025). Framework for Responsible Use of AI in Art Authentication. <https://itsartlaw.org/ai-and-art-authentication-guidelines/>, Accessed: 04.02.2026