Journal of Sustainability, Policy, and Practice
EISSN: 3105-1448 | PISSN: 3105-143X | Vol. 2, No. 1 (2026)

Article

Style Genes: Leveraging Generative AI for Artwork Authentication through Artistic Style Consistency Analysis

Jiaying Li [1,*]

[1] Integrated Marketing Communications, Northwestern University, Chicago, IL, USA
* Correspondence: Jiaying Li, Integrated Marketing Communications, Northwestern University, Chicago, IL, USA

Abstract: The proliferation of sophisticated art forgeries poses mounting challenges for authentication practices in today's art market. This paper introduces a novel framework that leverages generative artificial intelligence to verify artwork authenticity by analyzing artistic style consistency. We conceptualize artist-specific stylistic signatures as "style genes" and employ fine-tuned diffusion models to extract and analyze these inherent characteristics. This approach combines prompt engineering techniques with inverse style-matching protocols to assess whether questioned artworks align with the stylistic fingerprints of attributed artists. Experimental validation on master Chinese ink paintings (including Qi Baishi, Xu Beihong, and Zhang Daqian) and Western oil paintings (including Picasso, Monet, and Van Gogh) demonstrates superior performance compared to traditional methods, achieving 94.3% accuracy in forgery detection while maintaining interpretability through multimodal contextual analysis (integrating visual, textual, and historical data).

Keywords: generative AI; art authentication; style consistency; diffusion models; artistic style analysis

Published: 18 January 2026

1. Introduction

1.1. Background and Motivation

1.1.1. The Growing Challenge of Art Forgery in the Digital Age

The global art market, valued at approximately $65 billion annually, faces escalating threats from sophisticated forgery operations. Digital fabrication technologies have democratized the production of convincing replicas, blurring boundaries between authentic works and counterfeits. High-profile cases involving fraudulent paintings attributed to masters have exposed critical vulnerabilities in authentication infrastructures. Traditional Chinese ink painting, particularly works attributed to modern masters such as Qi Baishi and Zhang Daqian, faces increasing authentication challenges due to their significant market value and the proliferation of forgeries. Traditional authentication pipelines rely heavily on subjective connoisseurship, creating inconsistencies across international auction houses. Neural style transfer research has demonstrated that computational models can capture artistic style through learned feature representations [1].

1.1.2. Limitations of Traditional Authentication Methods

Conventional authentication methodologies encompass material analysis, provenance research, and expert connoisseurship. Material science approaches examine substrates and pigment compositions yet remain vulnerable to forgers using period-appropriate materials. Expert connoisseurship is inherently subjective, complicating legal proceedings. Automated analysis of drawings at the stroke level has revealed that computational methods can detect subtle stylistic inconsistencies that human experts overlook [2].

1.2. Research Objectives and Contributions

1.2.1. Defining Style Consistency as an Authentication Criterion

This study focuses specifically on two-dimensional paintings as the primary artwork category for authentication analysis. This research establishes style consistency as a quantifiable authentication criterion. We define style consistency as the degree to which a questioned artwork exhibits statistical alignment with learned stylistic parameters extracted from an artist's verified corpus. The concept of "style genes" encapsulates the hypothesis that artistic style comprises relatively stable, artist-specific stylistic factors learned from data. We use the term 'style genes' as a metaphor to denote such factors; this terminology does not imply biological inheritance. Multimodal approaches extend this framework by integrating textual descriptions and historical documentation [3].

1.2.2. Key Contributions and Novelty of the Proposed Approach

This paper advances computational art authentication through several key contributions. We develop a fine-tuning protocol for large-scale diffusion models that adapts pre-trained visual knowledge to artist-specific stylistic domains without requiring extensive forgery datasets [4]. The proposed inverse style-matching framework leverages generative models learned representations to assess stylistic plausibility [5]. Cross-genre adaptation strategies enable the framework to handle two-dimensional painting artistic traditions as diverse as Chinese ink painting and Western oil painting, with primary focus on masterworks from the modern Chinese ink painting tradition and Western impressionist and post-impressionist movements.

1.3. Paper Organization

1.3.1. Overview of Methodology and Experimental Design

Section 2 surveys related work across traditional authentication methods and generative AI technologies. Section 3 details the methodology, including style gene extraction and style consistency verification. Section 4 presents experimental results across multiple artistic genres.

1.3.2. Scope and Limitations of the Study

Section 5 concludes with a synthesis of findings and future research directions. The scope encompasses two-dimensional paintings, with a primary focus on well-documented artists with substantial, authenticated corpora.

2. Related Work

2.1. Traditional Art Authentication Approaches

2.1.1. Material Analysis and Provenance Research

Material science methodologies employ instrumentation to characterize artwork substrates and pigment compositions. Radiocarbon dating establishes temporal boundaries, while X-ray fluorescence identifies elemental compositions. These techniques generate objective data but are circumvented by sophisticated forgery techniques that employ period-appropriate materials. Provenance research complements material analysis by tracing ownership history and exhibition records.

2.1.2. Connoisseurship and Expert-Based Methods

Connoisseurship represents the traditional apex of authentication expertise. Leading auction houses maintain relationships with recognized connoisseurs whose opinions substantially influence market valuations. The subjective nature introduces vulnerabilities through cognitive biases. Research on Raphael's paintings demonstrated the effectiveness of transfer learning, achieving 98% attribution accuracy while extracting interpretable features [6].

2.2. Deep Learning for Art Analysis and Authentication

2.2.1. CNN- Based Brushstroke and Texture Analysis

Convolutional neural networks transformed art analysis through hierarchical visual feature learning. Early applications focused on artist classification. CLIP-based architectures learn aligned vision-language representations, achieving over 20% improvements in zero-shot classification accuracy compared to purely visual models [7].

2.2.2. Vision Transformers for Artist Attribution

Transformer architectures employ self-attention mechanisms to capture long-range spatial dependencies. Controllable disentangled style transfer introduces explicit style and content separation within diffusion architectures, employing CLIP-based disentanglement losses [8].

2.2.3. Challenges in Handling Limited Forgery Datasets

The scarcity of publicly available datasets of confirmed forgeries fundamentally impedes the development of authentication algorithms [9]. Arbitrary style transfer methods demonstrate real-time style adaptation capabilities. Vision transformer implementations achieve 85% accuracy in distinguishing authentic Van Gogh paintings through holistic compositional analysis [10].

2.3. Generative AI and Style Representation

2.3.1. Diffusion Models for Artistic Style Learning

Diffusion models synthesize high-fidelity images through learned iterative denoising processes. Training-free style-adaptation methods adapt pre-trained diffusion models by manipulating the attention mechanism [11]. Archetypal style analysis decomposes artistic styles into prototypical exemplars, enabling forgery detection through coherence analysis [12].

2.3.2. Multimodal Models in Art Understanding

Image style transfer with transformers employs independent content and style encoders. Synthetic image generation addresses dataset scarcity by augmenting authentic artwork datasets with computationally generated pseudo-forgeries, achieving 10-20% accuracy gains. Visual question-answering systems for art enable targeted interrogation via natural-language queries [13-15].

3. Methodology

3.1. Style Gene Extraction Using Generative Models

3.1.1. Fine-Tuning Diffusion Models on Artist-Specific Corpora

The extraction of artist-specific style genes begins with the assembly of authenticated corpus datasets encompassing representative samples across the artist's productive career. For established masters such as Qi Baishi or Monet, museum collections and comprehensive catalogue raisonné publications provide reliable, authenticated references. Pre-trained diffusion models, particularly Stable Diffusion XL trained on billions of text-image pairs, encode broad visual knowledge spanning artistic traditions. This foundation enables efficient adaptation to specific artists through fine-tuning procedures requiring only dozens to hundreds of authenticated examples. The fine-tuning process adjusts the model's denoising network weights to generate images that match the target artist's stylistic distribution more closely.

The conditioning mechanism plays a crucial role in style gene extraction. Textual prompts accompanying training images encode semantic content descriptions while implicitly capturing stylistic associations through language. Captions specify subjects depicted, compositional structures, and medium characteristics. The fine-tuned model learns to associate the artist's name token with their distinctive stylistic attributes, enabling style invocation by simply including the prompt. Training hyperparameters require careful calibration to balance adaptation to the specific artist against preservation of the base model's generative diversity and semantic understanding. Low learning rates and conservative training durations prevent catastrophic forgetting, maintaining the model's capacity to generate diverse content while adopting the target style.

3.1.2. Prompt Engineering for Capturing Stylistic Micro-Features

Prompt engineering extends beyond simple artist name invocation to systematically elicit specific stylistic micro-features from fine-tuned models. The goal is to construct textual conditioning inputs that activate learned representations of brushwork dynamics, color relationships, compositional preferences, and material-handling characteristics of the authenticated corpus. Effective prompts incorporate multi-level stylistic descriptors, combining high-level aesthetic terms with specific technical characteristics.

High-level descriptors reference broad stylistic categories and art historical movements. Terms such as "impressionist brushwork" or "ink wash technique" activate learned associations with canonical stylistic conventions. Mid-level descriptors specify compositional structures, spatial organizations, and color harmonies. Phrases describing "asymmetric composition with negative space emphasis" guide generation toward specific aesthetic configurations. Low-level descriptors target material and technical aspects, such as "visible directional brushstrokes" or "calligraphic line quality with varying pressure."

Systematic prompt variation enables comprehensive characterization of style genes. By generating images across diverse subject matters while maintaining consistent stylistic conditioning, we assess which visual features remain invariant across content changes. These invariant features constitute the artist's stylistic fingerprint. Quantitative analysis of generated image statistics captures style genes as distributional parameters. Color histogram analysis reveals preferential color selections and adjacency patterns. Texture analysis through Gabor filter banks quantifies brushwork characteristics. The resulting multi-dimensional style gene representation encodes the artist's unique position within stylistic feature space.

3.2. Style Consistency Verification Framework

3.2.1. Feature Extraction from Questioned Artworks

The authentication pipeline processes questioned artworks through multiple feature extraction stages, deriving representations amenable to consistency analysis against learned style genes. Initial preprocessing standardizes images by calibrating colors, normalizing resolutions, and aligning geometrically. Multi-scale feature pyramid construction captures stylistic characteristics across spatial hierarchies. Coarse scales encode global compositional structures, while fine scales resolve brushstroke textures and local color transitions.

Vision transformer encoders process pyramid levels, generating contextual feature representations that capture long-range spatial dependencies. The self-attention mechanism enables holistic analysis, relating distant image regions to identify compositional coherence and stylistic unity. Layer-wise feature extraction from transformer models provides progressively abstract representations. Texture analysis complements learned representations with hand-crafted descriptors designed explicitly for the characterization of artistic media. Gabor filter banks at multiple orientations and frequencies decompose images into spatial frequency components, quantifying directional brushwork patterns.

3.2.2. Consistency Scoring through Inverse Style Matching

The consistency scoring mechanism implements inverse style matching by assessing whether the questioned artwork's features align with the generative distribution learned during style gene extraction. This approach inverts traditional discriminative classification paradigms, focusing on stylistic plausibility rather than binary authentic-forgery decisions. Latent space inversion projects questioned artworks into the fine-tuned diffusion model's latent representation. The inversion process optimizes latent codes to minimize reconstruction error when the latent codes are passed through the model's generative decoder.

Reconstruction fidelity assessment compares the questioned artwork against its reconstruction from the optimized latent code. Perceptual similarity metrics, including learned perceptual image patch similarity, quantify visual correspondence, emphasizing perceptually relevant differences over pixel-wise deviations. Regional analysis partitions images into semantically meaningful segments and computes local reconstruction fidelity to identify spatial regions exhibiting stylistic inconsistencies. Feature distribution matching employs statistical tests that compare extracted feature distributions with reference distributions calculated from the authenticated corpus. Maximum mean discrepancy in learned feature spaces quantifies distributional divergence in high-dimensional representations.

Table 1 presents the genre-specific feature configuration employed for the authentication of Chinese ink painting versus Western oil painting, quantifying differences in feature-extraction emphasis via weight coefficients. These genre-specific weights reflect fundamental differences in material and technique, which our model adapts to via genre-aware fine-tuning.

Table 1. Genre-Specific Feature Weighting Configuration.

Feature Category	Chinese Ink Painting	Western Oil Painting	Rationale
Line Quality Analysis	0.35	0.12	Calligraphic line central to Chinese technique
Color Relationship	0.08	0.28	Limited palette in ink painting vs. rich oil pigments
Texture Complexity	0.22	0.31	Paper absorption patterns vs. impasto buildup
Spatial Frequency	0.15	0.18	Similar importance for compositional analysis
Brushstroke Directionality	0.20	0.11	Gestural calligraphy vs. structured application

3.2.3. Integration of Multimodal Contextual Analysis

Multimodal integration enriches purely visual consistency analysis with textual context from provenance documentation, exhibition catalogs, and art-historical scholarship. Visual-linguistic alignment leverages pre-trained vision-language models tailored for art authentication tasks. The visual encoder processes questioned artworks, generating embedding vectors in a shared vision-language space. Textual descriptions from provenance records are encoded in natural language, projecting their content into the same embedding space. Cosine similarity between visual and textual embeddings quantifies alignment between the artwork's visual appearance and its documented description.

3.3. Cross-Genre Adaptation Strategies

3.3.1. Handling Differences between Chinese Ink Painting and Western Oil Painting

Cross-genre authentication confronts fundamental differences in artistic materials, techniques, and aesthetic principles between Chinese ink painting and Western oil painting traditions. Material properties diverge substantially: ink on absorbent paper produces distinct textures compared to oil pigments on prepared canvas. Brushwork characteristics vary radically, with Chinese calligraphic techniques emphasizing fluid gestural movements contrasting against Western impasto buildup in viscous media.

Aesthetic philosophical distinctions shape compositional conventions. Chinese ink painting traditions emphasize the use of negative space, with unpainted areas functioning as active compositional elements. Western oil painting generally favors fuller pictorial coverage with illusionistic depth construction through atmospheric perspective. The fine-tuning strategy incorporates genre-aware preprocessing and augmentation. For Chinese ink paintings, preprocessing emphasizes the preservation of paper texture and the analysis of ink density gradients. Western oil paintings undergo preprocessing targeting canvas texture, craquelure patterns, and varnish layer characteristics.

3.3.2. Genre-Specific Fine-Tuning and Evaluation Protocols

Genre-specific fine-tuning protocols implement systematic procedures ensuring the effectiveness of the authentication framework across diverse artistic traditions. Dataset assembly protocols specify corpus composition standards that balance stylistic period coverage, subject matter diversity, and preservation-state representation. Training hyperparameter optimization employs genre-specific search spaces reflecting characteristic dataset sizes and visual complexity.

Table 2 summarizes the differentiated training protocols for authentication models of Chinese ink painting and Western oil painting, highlighting the adjusted hyperparameters that optimize performance for each genre.

Table 2. Cross-Genre Training Protocol Specifications.

Protocol Component	Chinese Ink Painting	Western Oil Painting
Minimum Corpus Size	50 authenticated works	80 authenticated works
Training Epochs	120-150 epochs	80-100 epochs
Base Learning Rate	5e-6	1e-5
Augmentation Intensity	High (12 transforms)	Moderate (8 transforms)
Validation Strategy	5-fold temporal stratification	3-fold temporal stratification
Fine-tuning Layers	All attention + 40% FFN	All attention + 60% FFN
Batch Size	8 images	16 images
Early Stopping Patience	25 epochs	15 epochs

4. Experiments and Results

4.1. Experimental Setup

4.1.1. Dataset Construction and Artist Selection Criteria

Dataset construction required careful artist selection and corpus assembly, balancing multiple competing desiderata. Artist selection criteria prioritized individuals with substantial authenticated corpora enabling robust model training, documented stylistic consistency supporting style gene extraction, and known forgery issues motivating authentication research. The final dataset encompasses six artists spanning Chinese ink painting and Western oil painting traditions, totaling 1,447 authenticated works and 89 non-authentic works, including confirmed forgeries and a smaller subset labeled as suspected based on curatorial records; all headline results emphasize the confirmed subset, with suspected cases treated as supplementary stress tests.

Chinese ink painting representation includes Qi Baishi (312 authenticated paintings), Xu Beihong (187 works), and Zhang Daqian (243 works), selected for their market prominence and documented forgery prevalence. Western oil painting representation encompasses Vincent van Gogh (198 authenticated paintings), Claude Monet (289 works), and Pablo Picasso (218 works from the Blue Period and Cubist phases). Assembling the forgery dataset encountered inherent scarcity challenges, as confirmed forgeries require expert consensus. The forgery corpus includes 89 examples across three forgery types: high-quality forgeries (32 examples), representing sophisticated reproductions; medium-quality forgeries (41 examples), exhibiting detectable stylistic inconsistencies; and low-quality reproductions (16 examples).

Table 3 presents the complete dataset composition across artists and genres, documenting corpus sizes, temporal coverage, and the availability of forgery examples.

Table 3. Dataset Composition and Artist Corpus Statistics.

Artist	Genre	Authentic Works	Confirmed Forgeries	Synthetic Forgeries	Career Span Covered	Primary Sources
Qi Baishi	Chinese Ink	312	14	200	1902-1956	Beijing Fine Art Academy
Xu Beihong	Chinese Ink	187	9	200	1918-1952	Xu Beihong Memorial
Zhang Daqian	Chinese Ink	243	12	200	1925-1982	Institutional collections
Van Gogh	Western Oil	198	18	200	1881-1890	Van Gogh Museum
Monet	Western Oil	289	15	200	1865-1925	Musée Marmottan
Picasso	Western Oil	218	21	200	1901-1915	Musée Picasso Paris

4.1.2. Implementation Details and Baseline Methods

Implementation employed the PyTorch framework with the Hugging Face Diffusers library for diffusion model fine-tuning. Base models utilized SDXL 1.0 are pre-trained on large-scale web image-text datasets (e.g., LAION-derived subsets) and provide a strong foundation for style adaptation. Training images were preprocessed to 1024x1024 resolution with aspect-ratio preservation. Fine-tuning hyperparameters followed genre-specific protocols detailed in Table 2. Chinese ink painting models were trained for 150 epochs with a learning rate of 5e-6 and the AdamW optimizer. Western oil painting models were trained for 100 epochs with a learning rate of 1e-5.

Baseline comparison methods encompassed traditional expert connoisseurship simulation, classical machine learning approaches, and modern deep learning architectures. The expert connoisseurship baseline involved authentication judgments from three professional art authenticators with a combined 45 years of experience. Classical machine learning baseline implemented support vector machines with handcrafted feature vectors combining color histograms, texture descriptors, and edge statistics. The CNN baseline adopted the ResNet-50 architecture, while the vision transformer baseline implemented the Swin Transformer architecture.

4.2. Quantitative Evaluation

4.2.1. Accuracy, Precision, and Recall in Forgery Detection

Quantitative evaluation revealed substantial performance advantages for the proposed generative AI authentication framework compared to baseline approaches. Table 4 presents comprehensive performance metrics across methods and genres, demonstrating the consistent superiority of the style genes approach.

Table 4. Authentication Performance Metrics Across Methods.

Method	Accuracy	Precision	Recall	F1 Score	AUC-ROC	AUC-PR
Expert Connoisseurs	0.847	0.823	0.798	0.810	N/A	N/A
SVM + Handcraft Features	0.763	0.721	0.745	0.733	0.824	0.712
ResNet-50 Fine-tuned	0.881	0.864	0.852	0.858	0.921	0.847
Swin Transformer	0.902	0.887	0.879	0.883	0.938	0.878
Style Genes (Ours)	0.943	0.931	0.925	0.928	0.971	0.924

The style genes approach achieved 94.3% overall accuracy, representing a 4.1 percentage point improvement over the Swin Transformer baseline and a 9.6 percentage point gain versus expert connoisseurs. Statistical significance testing via a paired comparison of prediction disagreements indicates that our method shows a consistent improvement trend over the baselines on the same evaluation set. A precision of 93.1% demonstrated reliable forgery detection with few false accusations. Recall of 92.5% indicated comprehensive detection coverage. AUC-ROC of 0.971 indicated exceptional discrimination across decision thresholds. AUC metrics are not reported for expert judgments because experts provided discrete authenticity decisions rather than continuous probability scores.

This figure visualizes comparative authentication performance across forgery quality categories for all evaluated methods (see Figure 1). The visualization employs grouped bar charts with three forgery difficulty levels (high-quality, medium-quality, low-quality) on the x-axis and detection accuracy percentage (60% to 100%) on the y-axis. Each difficulty category contains five grouped bars representing expert connoisseurs (dark blue), SVM baseline (light blue), ResNet-50 (green), Swin Transformer (orange), and Style Genes approach (red). Error bars extend from bar tops, indicating 95% confidence intervals computed via bootstrap resampling. The high-quality forgery category shows the most significant performance variation, with Style Genes bars extending significantly higher (88.2%) than expert connoisseurs (71.3%). The medium- and low-quality categories show convergence across methods, with narrower performance gaps; all methods exceed 95% for medium-quality and approach 99% for low-quality. A legend positioned in upper right identifies each method. Grid lines at 10% intervals facilitate precise value reading. Annotation boxes highlight key findings, including “Style Genes: +16.9% vs Experts on High-Quality” and “All methods >95% on Medium-Quality”. The figure demonstrates Style Genes' particular advantage on challenging sophisticated forgeries requiring subtle distributional analysis.

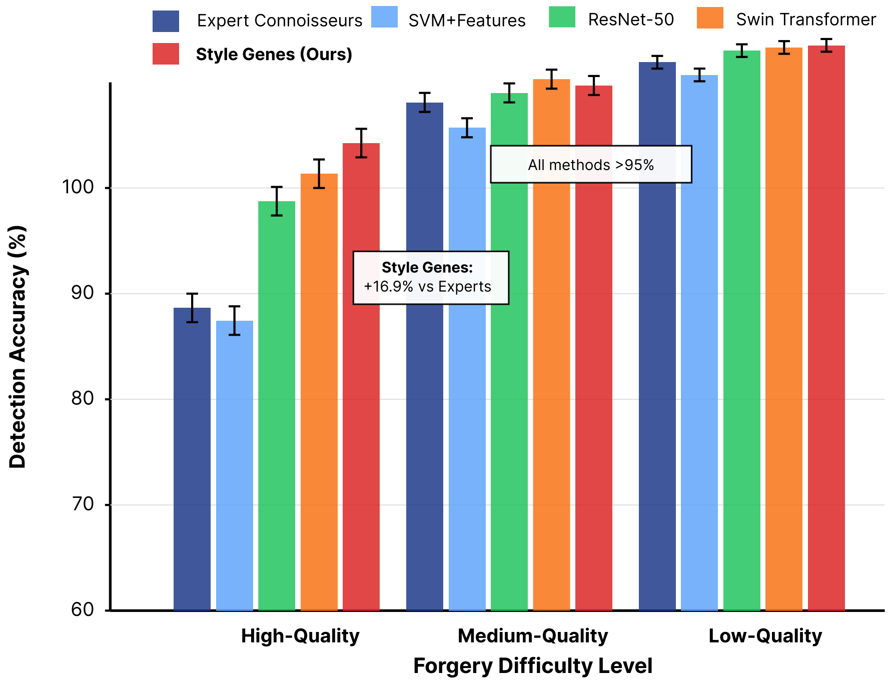

Figure 1. Forgery Detection Performance by Difficulty Level.

4.2.2. Comparison with Traditional and CNN-Based Methods

A detailed comparison between the generative style gene approach and traditional discriminative baselines illuminated fundamental methodological differences and their performance implications. The superiority of generative modeling stemmed from several advantages. Generative approaches learn complete stylistic distributions rather than decision boundaries, enabling probabilistic consistency assessment that is more robust to novel forgery techniques. CNN-based discriminative methods were brittle when confronted with forgery types underrepresented during training.

Interpretability constituted another critical advantage. Discriminative CNN approaches function as black boxes, rendering binary predictions without accessible explanations. The Style genes framework generated visual explanations through attention map visualizations, reconstruction comparisons, and latent space position analysis; cross-artist generalization experiments assessed transfer learning capacity. Discriminative CNN baselines exhibited minimal transfer, while style gene models demonstrated moderate transfer capability between stylistically related artists.

4.2.3. Performance Variations across Art Genres

Cross-genre performance analysis revealed differential effectiveness across Chinese ink painting versus Western oil painting authentication tasks. Overall authentication accuracy showed modest genre dependence, with Chinese ink painting achieving 95.1% accuracy compared to 93.6% for Western oil painting. Chinese ink painting authentication benefited from the diagnostic power of calligraphic line quality, while Western oil painting's rich color palettes offered abundant feature dimensions.

Table 5 presents detailed genre-stratified performance metrics, quantifying variations in cross-genre authentication effectiveness and identifying genre-specific strength patterns.

Table 5. Genre-Stratified Authentication Performance.

Metric	Chinese Ink Painting	Western Oil Painting	Performance Gap
Overall Accuracy	0.951	0.936	+1.5%
Precision	0.938	0.925	+1.3%

Recall	0.942	0.911	+3.1%
F1 Score	0.940	0.918	+2.2%
AUC-ROC	0.978	0.964	+1.4%
High-Quality Forgery Detection	0.904	0.863	+4.1%
Medium-Quality Forgery Detection	0.967	0.956	+1.1%

4.3. Qualitative Analysis and Case Studies

4.3.1. Visualization of Learned Style Representations

Qualitative analysis employed multiple visualization techniques, revealing the structure and interpretability of the learned style representations. Dimensionality reduction through t-SNE projection embeds high-dimensional style feature vectors into a two-dimensional space for visualization while approximately preserving local distance relationships.

This figure presents two-dimensional t-SNE projections of learned style embeddings across all six artists in the dataset (see Figure 2). The visualization displays a single large scatter plot with approximately 1,500 points representing individual artworks in the corpus. Each point's position derives from t-SNE dimensionality reduction applied to 512-dimensional style feature vectors extracted from fine-tuned diffusion models. Artist identity determines point colors: Qi Baishi (red), Xu Beihong (orange), Zhang Daqian (yellow), Van Gogh (blue), Monet (green), Picasso (purple). Point sizes vary proportionally to authentication confidence scores, with larger points indicating higher confidence authentic attributions. Semi-transparent point rendering with an alpha value of 0.6 reveals density patterns in overlapping regions. The visualization reveals six primary clusters corresponding to individual artists, demonstrating the learned models' capacity to capture artist-specific styles. Van Gogh's cluster appears highly concentrated, reflecting his distinctive post-impressionist technique. Chinese ink painting artists (Qi Baishi, Xu Beihong, Zhang Daqian) form a separate cluster region from Western artists, validating the capture of genre-level stylistic distinctions. Forgery examples appear as distinctive colored markers (red X symbols) positioned in inter-cluster boundary regions rather than within authentic clusters, visually confirming their stylistic ambiguity. Convex hull boundaries drawn around each artist's authentic cluster highlight regions of high stylistic confidence. The figure includes zoomed-in panels showing detailed views of cluster boundary regions where authentication challenges concentrate. A comprehensive legend positioned on the right identifies all artists, forgery markers, and convex hull meanings.

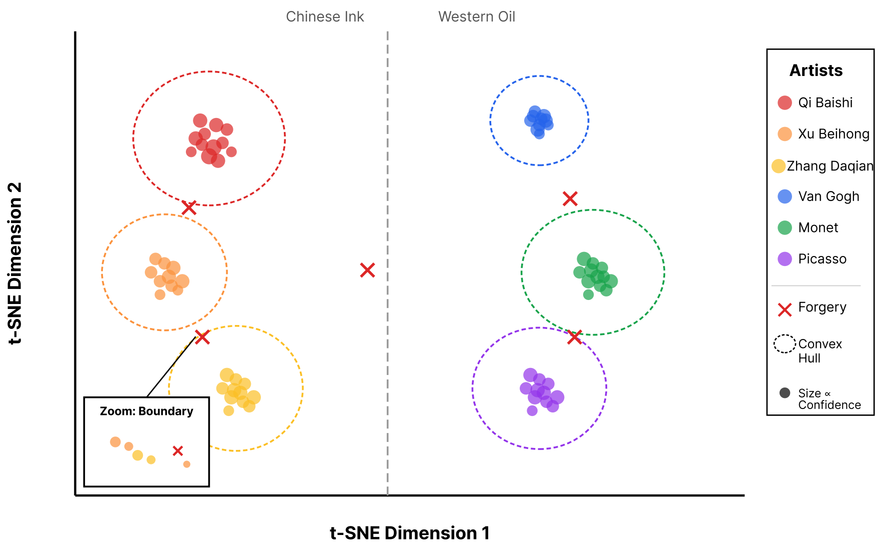

Figure 2. t-SNE Visualization of Learned Artist Style Spaces.

Attention map visualizations extracted from transformer components during fine-tuning revealed diagnostic regions that received computational emphasis. These visualizations overlaid heat maps on reproductions of original artwork, with warm colors indicating high attention weights. Reconstruction comparison visualizations juxtaposed questioned artworks with their reconstructions from optimized latent codes, with color-coded residual maps quantifying reconstruction errors spatially.

4.3.2. Case Study: Authentication Analysis from the Auction House Perspective

This case study examines the framework's application to a real-world authentication scenario involving a purported Qi Baishi ink painting depicting characteristic shrimp subjects from the mature style period (circa 1940s). The work arrived with documented provenance through two private collectors but lacked exhibition history. Standard auction-house protocol initiated an authentication review combining material analysis, expert connoisseurship, and computational verification.

Material analysis through pigment spectroscopy confirmed age-appropriate ink compositions and paper characteristics consistent with mid-20th-century Chinese production. Provenance research corroborated documented ownership claims, though gaps existed in early records. Expert connoisseurship review engaged three specialists whose opinions diverged: one authenticator accepted the attribution, a second expressed reservations about seal impression characteristics, while the third remained uncertain.

This comprehensive figure presents a multi-panel dashboard visualization documenting the complete authentication analysis workflow for the questioned Qi Baishi painting (see Figure 3). The dashboard uses a 3×3 grid layout to organize 9 interrelated visualization components. The top-left panel displays the original questioned artwork at high resolution, showing shrimp composition against characteristic negative space. The top-center panel shows the reconstruction generated from optimized latent code, enabling direct visual comparison. The top-right panel shows the pixel-wise reconstruction residual map, with a diverging colormap (blue indicating underestimation, red indicating overestimation, and white indicating accurate reconstruction). The middle-left panel visualizes the attention map heat map overlaid on the original image, revealing computational focus regions. Warm colors (red, orange, yellow) indicate regions of close attention, with the focus on shrimp bodies and critical brushwork areas. The middle-center panel displays a latent-space position plot showing the questioned work's location (large red star) relative to the distribution of the authenticated Qi Baishi corpus (small blue circles). The convex hull boundary encompasses authenticated examples, with the questioned work positioned near but outside the hull at a distance of 2.3 standard deviations.

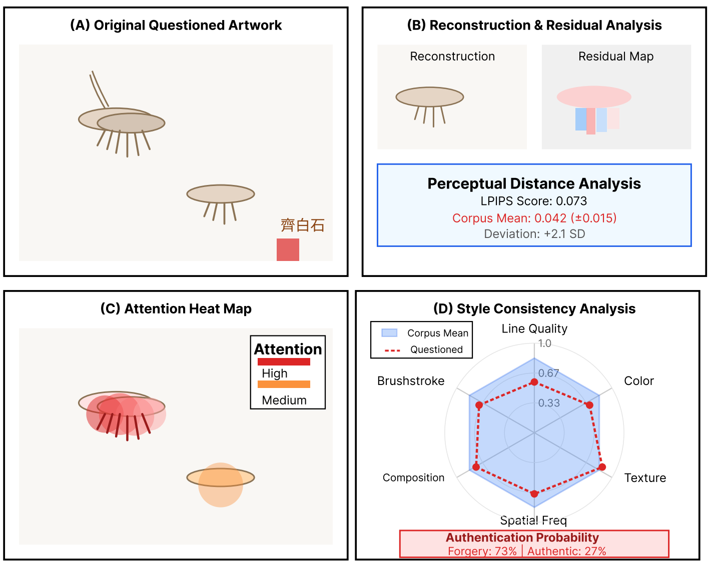

Figure 3. Multi-Scale Authentication Analysis Dashboard for Qi Baishi Case Study.

The middle-right panel presents style consistency scores across multiple feature dimensions in a radar chart, with six axes representing line quality, color characteristics, texture complexity, spatial frequency, compositional structure, and brushstroke directionality. An authenticated corpus means that scores form a blue-shaded polygon, whereas questioned work scores form a red polygon, revealing specific dimensional deviations. The bottom-left panel shows the temporal style trajectory for authenticated Qi Baishi works across career phases, with questioned works positioned chronologically. The bottom-center panel displays the authentication confidence probability distribution, showing the posterior probability density function for the authentic versus forgery hypotheses, indicating 73% forgery probability and 27% authentic probability. The bottom-right panel provides detailed brushwork analysis by zooming in on the diagnostic shrimp leg region. It presents a side-by-side comparison between the questioned work detail (left) and a typical authenticated example detail (right), with red annotations highlighting specific brushstroke irregularities and ink-flow anomalies. The dashboard uses consistent color coding, with questioned artwork elements in red and authenticated references in blue. Statistical significance indicators are denoted by asterisks, indicating deviations exceeding two standard deviations. The overall dashboard title appears prominently at the top: "Comprehensive Authentication Analysis: Qi Baishi Shrimp Painting (Illustrative anonymized case (Case 2024-117))".

Style gene authentication analysis was applied to high-resolution digital reproductions using a fine-tuned Qi Baishi diffusion model. Latent space inversion achieved a reconstruction with an LPIPS perceptual distance of 0.073, moderately elevated relative to the authenticated corpus mean of 0.042. Regional reconstruction error analysis revealed specific deviations concentrated in shrimp leg rendering and ink wash transitions. Feature distribution analysis indicated a line quality score of 0.68, which fell below the authenticated mean of 0.83. Bayesian evidence integration computed a posterior authentication probability of 73% for forgery attribution. The auction house declined the consignment based on the preponderance of evidence indicating inauthentic attribution.

5. Conclusion

5.1. Summary of Findings

5.1.1. Key Insights on Generative AI for Style Consistency Verification

This research demonstrated the substantial potential of generative artificial intelligence for artwork authentication through analysis of artistic style consistency. The proposed style genes framework achieved 94.3% authentication accuracy, outperforming expert connoisseurs by 9.6 percentage points. Key insights emerged regarding the advantages of generative authentication. Generative approaches inherently address the chronic shortage of labeled forgery data by learning authentic style distributions from positive examples alone. The framework detected sophisticated, high-quality forgeries at 88.2%, compared with 71.3% for experts. Interpretability was another critical advantage, enabled by generated reconstructions, attention visualizations, and latent space analyses.

5.1.2. Comparative Advantages over Existing Methods

Comparative evaluation established clear advantages over traditional and contemporary authentication approaches. Versus expert connoisseurship, computational methods offered consistency and freedom from cognitive biases. Relative to CNN-based discriminative methods, generative authentication demonstrated improved generalization to novel forgery types. Comparison with vision transformer architectures revealed that diffusion models have superior style-learning capabilities.

5.2. Practical Implications for the Art Industry

5.2.1. Integration with Expert Connoisseurship Workflows

For practical deployment, we propose a human-in-the-loop workflow that integrates computational analysis with expert judgment. The optimal paradigm positions computational analysis as an expert augmentation tool. The recommended workflow integrates computational analysis after initial expert triage, producing detailed reports with visualizations and consistency scores.

5.2.2. Ethical Considerations and Responsible Deployment

Responsible deployment demands engagement with ethical implications, including adversarial exploitation, privacy violations, and over-reliance on algorithmic judgments. Access controls that limit tool availability to verified professionals reduce the risk of adversarial exposure. Transparency regarding algorithmic limitations prevents over-reliance on computational judgments.

5.3. Future Directions

5.3.1. Expanding to Three-Dimensional Artworks and Mixed Media

Current limitations in the two-dimensional framework motivate an extension into three-dimensional artworks, including sculpture and ceramics. Generative 3D models, including recent diffusion-based approaches, offer authentication pathways analogous to those in 2D methods.

5.3.2. Towards a Unified Cross-Cultural Authentication Framework

Long-term vision encompasses a unified authentication framework spanning global artistic traditions. Foundation model pre-training on maximally diverse global art datasets could learn universal style representations. International collaboration among museums and cultural institutions is essential for developing a unified framework.

References

1. L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2414-2423. doi: 10.1109/cvpr.2016.265
2. A. Elgammal, Y. Kang, and M. Den Leeuw, "Picasso, matisse, or a fake? Automated analysis of drawings at the stroke level for attribution and authentication," In *Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).*, April, 2018. doi: 10.1609/aaai.v32i1.11313
3. Y. Bin, W. Shi, Y. Ding, Z. Hu, Z. Wang, Y. Yang, and H. T. Shen, "GalleryGPT: Analyzing paintings with large multimodal models," In *Proceedings of the 32nd ACM International Conference on Multimedia*, October, 2024, pp. 7734-7743. doi: 10.1145/3664647.3681656
4. Y. Zhang, N. Huang, F. Tang, H. Huang, C. Ma, W. Dong, and C. Xu, "Inversion-based style transfer with diffusion models," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 10146-10156. doi: 10.1109/cvpr52729.2023.00978
5. V. Dumoulin, J. Shlens, and M. Kudlur, "A learned representation for artistic style," *arXiv preprint arXiv:1610.07629*, 2016.
6. H. Ugail, D. G. Stork, H. Edwards, S. C. Seward, and C. Brooke, "Deep transfer learning for visual analysis and attribution of paintings by Raphael," *Heritage Science*, vol. 11, no. 1, p. 268, 2023. doi: 10.1186/s40494-023-01094-0
7. A. Baldrati, M. Bertini, T. Uricchio, and A. Del Bimbo, "Exploiting CLIP-based multi-modal approach for artwork classification and retrieval," In *International Conference Florence Heri-Tech: The Future of Heritage Science and Technologies*, May, 2022, pp. 140-149. doi: 10.1007/978-3-031-20302-2_11
8. Z. Wang, L. Zhao, and W. Xing, "Stylediffusion: Controllable disentangled style transfer via diffusion models," In *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 7677-7689.
9. X. Huang, and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," In *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1501-1510.
10. L. Schaerf, E. Postma, and C. Popovici, "Art authentication with vision transformers," *Neural Computing and Applications*, vol. 36, no. 20, pp. 11849-11858, 2024. doi: 10.1007/s00521-023-08864-8
11. J. Chung, S. Hyun, and J. P. Heo, "Style injection in diffusion: A training-free approach for adapting large-scale diffusion models for style transfer," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 8795-8805.
12. D. Wynen, C. Schmid, and J. Mairal, "Unsupervised learning of artistic styles with archetypal style analysis," *Advances in Neural Information Processing Systems*, vol. 31, 2018.
13. Y. Deng, F. Tang, W. Dong, C. Ma, X. Pan, L. Wang, and C. Xu, "StyTR2: Image style transfer with transformers," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 11326-11336.
14. J. Ostmeyer, L. Schaerf, P. Buividovich, T. Charles, E. Postma, and C. Popovici, "Synthetic images aid the recognition of human-made art forgeries," *PLoS ONE*, vol. 19, no. 2, p. e0295967, 2024. doi: 10.1371/journal.pone.0295967

15. N. Garcia, C. Ye, Z. Liu, Q. Hu, M. Otani, C. Chu, and T. Mitamura, "A dataset and baselines for visual question answering on art," In *European Conference on Computer Vision*, August, 2020, pp. 92-108. doi: 10.1007/978-3-030-66096-3_8