

<https://doi.org/10.1038/s40494-026-02304-1>

Toward enhanced unsupervised clustering of 20th century Korean paintings via multimodal features

Check for updates

Seohyun Baek¹, So-Jeong Park², So-Eun Park², You-Min Im³, Jongwon Choi² & Bo-A Rhee^{2,3}✉

This study presents a machine learning framework for analyzing and clustering modern and contemporary Korean paintings based on image data. A pretrained vision–language architecture combined with multi-layered analysis was used to efficiently extract detailed formal characteristics, including color features from multiple spaces and quantified texture. The extracted feature vectors are clustered and evaluated under majority-label assignment, achieving 82.4% overall accuracy, outperforming single-feature baselines (RGB 82.0%, HSV 81.3%, histogram 51.0%, LBP 68.8%, and GLCM 73.7%). The proposed method achieves higher per-artist precision, better boundary-case discrimination, and greater robustness for low-sample categories. Representative images from artist clusters encapsulate unique color and texture. Analysis was extended using automatic image captioning and zero-shot style assignment via matching image and text embeddings. The findings demonstrate that machine learning–based image analysis provides an effective and objective methodology for identifying and distinguishing visual characteristics of modern and contemporary Korean paintings, offering a quantitative approach to art-historical interpretation.

Analyzing artworks is inherently complex^{1,2}. Art experts describe visual features such as space, texture, edges, form, shape, color, composition, lighting, brushstroke, tone, and line^{3,4}. They also evaluate movement, harmony, balance, contrast, proportion, and pattern⁵. Collectively, these formal characteristics provide quantifiable data that support the systematic study of an artist's techniques, intentions, and narrative structures.

Each artwork carries a distinctive visual signature⁶. This stylistic uniqueness enables the identification of artistic relationships and contributes to understanding broader art movements and genres⁷. Quantitative analysis of brushstroke configurations, for example, serves as a robust indicator of an artist's stylistic tendencies⁸. However, defining such styles remains challenging because stylistic boundaries frequently overlap, and artists may adopt multiple approaches over time, complicating recognition and classification^{9,10}. For instance, Pablo Picasso's body of work encompasses both Surrealist and Cubist idioms, illustrating continuous stylistic evolution. Human judgment in art is also inherently subjective and often contentious^{11,12}, as expert evaluations rely heavily on individual experience and disciplinary knowledge¹³.

The intersection of art and technology has a long history, with scientific techniques assisting art experts since the early eighteenth century¹⁴. The development of computerized imaging methods—including ultraviolet fluorescence, infrared reflectography, stereomicroscopy, and

X-radiography¹⁵—together with advances in machine learning and computer vision algorithms¹⁶, has enabled computational analysis to become an essential interdisciplinary tool for examining artworks and refining our understanding of their material and visual properties^{17,18}. The increasing accessibility of large-scale image datasets, such as WikiArt and ImageNet¹⁹, has further expanded the scope and accuracy of computational art analysis^{4,20}.

Computational algorithms have proven valuable for identifying and comparing artistic styles and for quantifying similarities among paintings. Machine learning models can encode discriminative visual features²¹ and serve as powerful tools for detecting forgeries and authenticating uncertain artworks^{1,21,22}. Deep neural networks further reveal latent visual patterns, stylistic signatures, and semantic relationships across artworks while performing style classification^{9,12}. Empirical evidence suggests that these computational approaches frequently outperform even highly trained experts in accuracy²³, reinforcing computational aesthetics as a rigorous and reproducible framework for analyzing visual properties within the expanding field of digital art history^{24,25}.

Research on machine learning–based art classification has employed a variety of approaches, primarily utilizing Convolutional Neural Networks (CNNs)^{26–28} and Attention Mechanisms^{29–31}. Traditional machine learning techniques have also been applied, combining feature extraction methods

¹Content Research Division, Electronics and Telecommunications Research Institute, Daejeon, Korea. ²Department of Advanced Imaging, GSAM, Chung-Ang University, Seoul, Korea. ³School of Art & Technology, Chung-Ang University, Anseong, Gyeonggi-do, Korea. ✉ e-mail: boa.rhee@gmail.com

such as Principal Component Analysis (PCA)³² and Histogram of Oriented Gradients (HOG)³³ with K-means clustering^{34–36}. However, the majority of these studies have focused predominantly on Western artworks³⁷, leading to a systematic bias toward specific artistic movements and cultural contexts. In contrast, research on traditional Korean paintings and East Asian visual heritage remains relatively limited.

For example, Elgammal et al. demonstrated that unsupervised clustering applied to large-scale Western painting datasets can successfully recover canonical art-historical movements and even reveal unexpected affinities between artists. Moreover, CNNs trained solely on stylistic labels were able to learn temporally coherent progressions without explicit chronological metadata^{21,38,39}. In a complementary study, Kim et al. explored content-based clustering, showing that subject detectors achieved high precision in identifying dominant categories, while co-occurrence analysis produced semantically coherent thematic networks⁴⁰.

Methods

Classification of Art-historical movements

A substantial body of research in computational art analysis has focused on the automatic classification of artworks according to categories such as artist, style, or genre. Several studies have specifically addressed automatic artist classification and identification^{1,10}, style classification^{5,9}, and genre classification^{17,41}. Previous research on artist identification, artistic style recognition, and art movement classification has extracted a wide range of features, encompassing both handcrafted features for traditional machine learning approaches^{6,42} and deep features representing digitized paintings for deep learning models^{1,10,43}.

Regarding handcrafted features, researchers have employed both low-level and high-level descriptors. Low-level features often include color and texture attributes^{2,44,45}, whereas high-level semantic features capture compositional or contextual information⁴⁶. Some studies have also combined both types of features to enhance classification accuracy⁵. In the case of Vincent van Gogh, for example, quantitative descriptors such as brush-stroke distribution, orientation, width, length, color palette, composition, and shape have been instrumental in defining the artist's distinctive stylistic signature and facilitating artist identification^{14,44,47}.

Ahmed Elgammal and collaborators have further advanced this field by applying clustering techniques to group artists and artistic schools based on stylistic similarities derived from computational analyses. Their research illustrates how visual features can be quantitatively measured and interpreted to reveal meaningful structural relationships within art history. For instance, large-scale datasets of paintings analyzed through unsupervised clustering have successfully grouped artists not only by individual style but also according to broader movements such as Impressionism, Cubism, and Abstract Expressionism^{38,39}. The results demonstrated that computationally derived clusters frequently aligned with established art-historical taxonomies, while simultaneously uncovering unexpected affinities among artists traditionally regarded as distinct. Such findings underscore the potential of machine learning to both validate existing art-historical frameworks and uncover latent stylistic connections that enrich our understanding of artistic evolution.

In addition to studies centered on feature-based style classification, such as *The Shape of Art History in the Eyes of the Machine*²¹, scholars have also explored content-based clustering approaches. Elgammal et al. demonstrated that convolutional neural networks trained solely on style labels can internally organize artworks into a continuous and historically coherent temporal sequence, thereby reconstructing stylistic movements without explicit temporal or contextual input. Their findings further indicated that the learned visual factors correspond closely to canonical art-historical distinctions, such as those proposed by Heinrich Wölfflin, and that certain artists emerge as prototypical exemplars positioned at stylistic extremes.

In *Computational Analysis of Content in Fine Art Paintings*⁴⁰, the analytical focus shifts from style to content. This work examines the presence and distribution of objects and subjects within paintings, the reliability

of automated content detection—achieving approximately 68% precision for dominant categories—and the co-occurrence patterns among subjects that define semantically meaningful relationships between content types.

Collectively, these studies demonstrate that machine learning can support clustering along multiple dimensions—both stylistic (temporal or school-based) and thematic (content-based)—thereby enriching art-historical analysis through the revelation of formal stylistic trajectories as well as interconnected thematic networks.

Focusing on classification models

Recent studies on artwork classification have primarily employed either Generative or Discriminative models, or combinations of both^{30,48,49}. Some approaches rely exclusively on Generative models, while others focus on Discriminative frameworks. In addition, Graphical models have been investigated for representation learning, and Hypothesis Matching models have been applied for comparative data analysis.

The classification of artworks using Convolutional Neural Networks (CNNs) depends on the extraction of discriminative visual features while maintaining computational efficiency. Recent advancements in Attention Mechanisms have enhanced the ability of models to capture intricate stylistic details and to distinguish subtle variations between art styles, thereby addressing the limitations of conventional CNN architectures^{30,48,49}. By selectively emphasizing semantically relevant image regions, attention-based approaches substantially improve classification accuracy.

The extraction of color and texture features remains fundamental for differentiating artistic styles, as distinct movements often display characteristic color palettes and brushwork patterns. Integrating these low-level features within neural network architectures improves learning efficiency and classification accuracy in automated art analysis^{50–52}. The Multi-Class Kernel Method further enhances robustness, particularly in the classification of figurative styles that incorporate diverse feature components such as chromatic attributes, texture, morphology, and compositional structure. By mapping these features into high-dimensional spaces, the method captures complex non-linear relationships, thereby enabling more precise differentiation across artistic styles⁵³.

Transfer learning has also become a critical approach, leveraging pre-trained models to enhance performance on new tasks through the reuse of knowledge from large-scale datasets. This process reduces the need for extensive computational resources while maintaining high classification accuracy⁴⁸.

Large-scale open datasets have played a central role in advancing computational art analysis. WikiArt⁵⁴ remains one of the most comprehensive, containing approximately 150,000 artworks from 2,500 artists^{8,10}. Other frequently used datasets include ArtCyclopedia⁵⁵, Artstor Digital Library⁴⁶, BBC Painting Dataset⁵⁶, Mark Harden's Archive⁵, ABC Gallery⁹, and Artlex & CARLI Digital Collections⁷. Additional image data have been collected from open-access platforms such as Wikipedia, Flickr, and online museum archives^{57,58}.

Park et al.⁵⁹ selected images from 25 artists within the WikiArt dataset and addressed class imbalance using a weighted cross-entropy loss function. Dataset expansion was achieved through augmentation techniques, including resizing, horizontal flipping, and rotation using the *Albumentations* library. Contrast Limited Adaptive Histogram Equalization (CLAHE) was employed to improve contrast, and the CutMix algorithm was applied to enhance texture representation. By fine-tuning a ResNet50 model—adjusting fully connected layers and freezing convolutional weights—the authors achieved significant improvements in artist classification accuracy.

Previous research has explored the computational classification of traditional Chinese paintings (TCP)^{60–62}. Li and Wang⁶⁰ utilized wavelet transforms and two-dimensional Multi-Resolution Hidden Markov Models (MHMMs) to categorize Chinese ink paintings according to style and artist. Jiang et al.⁶¹ distinguished TCP images from non-TCP artworks and further classified them into Gongbi (meticulous brushwork, 1,889 images) and Xieyi (freehand) styles using low-level features—such as color, texture, and edge characteristics—combined with a hybrid classifier integrating decision

trees and Support Vector Machines (SVM). Their approach achieved practically viable accuracy for the differentiation of traditional painting styles. Lu et al.⁶² developed a TCP classification framework encompassing four artistic movements (Xieyi, Gongbi, Goule, and Shese) and six painters, employing Bayesian classifiers, k-Nearest Neighbor (k-NN), fuzzy C-means clustering, and non-linear multi-class SVMs to compare classification performance across techniques.

More recent studies have increasingly emphasized improving classification accuracy through data augmentation and the use of advanced deep learning models. Baldrati et al.⁶³ introduced a CLIP-based multimodal framework that combines textual and visual features using the NoisyArt dataset, thereby enhancing both classification and retrieval tasks. Their results demonstrated the effectiveness of multimodal learning in computational art analysis, highlighting the value of cross-modal feature integration. Zhong et al.⁶⁴ proposed a Two-Channel Dual-Path Network (FPTD) incorporating RGB and brushstroke texture information to improve the fine-art painting classification process. This method employed a Gray-Level Co-Occurrence Matrix (GLCM) to extract texture features from multiple directions, achieving more accurate classification of style, artist, and genre while improving model generalization.

Kim et al.⁶⁵ advanced this line of inquiry by developing a proxy-learning approach that integrates pre-trained language models with visual data for artistic style analysis. By modeling the semantic relationships between textual descriptions and visual features, their method extracts meaningful visual concepts that significantly enhance automated artwork classification and interpretation. This interdisciplinary approach extends conventional feature extraction methodologies and provides new perspectives on multimodal computational analysis in art history.

Modern and contemporary Korean artists

Prominent figures in modern and contemporary Korean art include Kim Ki-chang, Kim Whan-ki, Do Sang-bong, Park Soo-keun, Yoo Young-kuk, Lee Jung-seob, Chun Kyung-ja, Chang Uc-chin, Byun Kwan-sik, Lee Sang-beom, and Byun Jong-ha. Active throughout the twentieth century, these artists played a pivotal role in defining Korean modernism and shaping a distinct national artistic identity within the global art world. Against the backdrop of Korea's turbulent modern history, they cultivated highly individual artistic vocabularies that collectively illustrate the evolution of Korean art. Among them, Kim Whan-ki, Park Soo-keun, Yoo Young-kuk, Lee Jung-seob, and Chang Uc-chin are widely recognized as the five leading second-generation Western-style painters and the first generation of Korean modernists⁶⁶.

Kim Whan-ki (1913–1974) and Yoo Young-kuk (1916–2002) explored Korean modernity through experimental abstraction that bridged traditional aesthetics and contemporary expression. Yoo Young-kuk pursued geometric purity in abstraction, extending pre-war modernist formalism⁶⁷, whereas Kim Whan-ki incorporated familiar motifs from everyday life to articulate the spiritual dimension of modern Korean experience⁶⁸. Lee Jung-seob (1916–1956), characterized by a restrained palette and dynamic brushwork, depicted local emotions and resilience through recurring motifs of cows and children⁶⁹. Chang Uc-chin (1917–1990) engaged with formative compositions inspired by rural life and childhood memories, while Park Soo-keun (1914–1965) portrayed the dignity of ordinary people during the Japanese colonial period and the Korean War, capturing their perseverance and humanity with what critics describe as a “sincere heart and gentle gaze”⁷⁰.

Except for Park Soo-keun, the other four artists were members of the New Realism Group (Shinsasilpa), an influential collective founded in July 1947⁷¹. Comprising mainly graduates of Japanese art academies, the group sought to synthesize modernist aesthetics with a renewed Korean identity, responding to the cultural and political upheavals of the post-liberation period.

Kim Ki-chang (1914–2001), an Oriental painter, modernized *Joseon* era folk and genre painting through his distinctive “Foolish Painting Style,” reinterpreting traditional narratives with contemporary sensibilities⁷². Do

Sang-bong (1902–1977) infused Korean sentiment into realist painting from the late 1920s to the 1970s, characterized by balanced composition and rich chromatic ton⁶⁶. Chun Kyung-ja (1924–2015), the only female artist among the eleven, established a unique aesthetic by fusing traditional Korean color palettes with bold, expressive hues. During a period dominated by monochrome ink painting, she introduced new possibilities for chromatic painting, integrating Korean emotionality with modern aesthetics⁷³.

Byun Kwan-sik (1899–1976) preserved the essence of Korean ink traditions while pioneering the “Sojeong Style,” distinguished by diverse ink techniques and innovative compositional layouts⁷⁴. Lee Sang-beom (1897–1972), influenced by photography and Western painting since the 1920s, became a master of ink-wash landscapes. Through the development of the “Cheongjeon Style,” he transformed conceptual landscapes into realistic depictions of nature, employing *mijeomjun* techniques to modernize traditional *sansu* painting⁷⁵. Byun Jong-ha (1926–2000) expanded the spatial vocabulary of Korean painting through his “three-dimensional painting” style, developed during the 1960s. By introducing depth and materiality to pictorial surfaces, his work blurred the boundaries between painting and sculpture, paving new directions in Korean modern art⁷⁶.

Dataset composition

In this study, we constructed a dataset comprising 1,100 paintings from eleven representative modern and contemporary Korean artists, with 100 works per artist (Table 1). The selected artists include Kim Ki-chang, Kim Whan-ki, Do Sang-bong, Park Soo-keun, Yoo Young-kuk, Lee Sang-beom, Lee Jung-seob, Chun Kyung-ja, Chang Uc-chin, Byun Kwan-sik, and Byun Jong-ha. The research team conducted a rigorous selection process to ensure the dataset's representativeness in terms of both stylistic diversity and historical significance within twentieth-century Korean art.

The majority of images were sourced from the National Museum of Modern and Contemporary Art (MMCA), which, as of 28 October 2024, provides digital access to 11,479 images, including 3608 modern paintings. To further expand the dataset, we collected high-resolution images from publicly accessible and reputable platforms such as Google Arts & Culture and Google Search. Authenticity and data integrity were prioritized by exclusively selecting artworks published on authoritative sources that provide verified metadata, including title, creation year, and medium.

For the acquisition of supplementary images, Google Search was used with the artist's full name as a keyword and the “large size” filter enabled to obtain high-resolution results. Missing metadata (e.g., title, year of creation) were recovered through reverse image searches using Google Lens, cross-referenced with artist foundation archives and other credible institutional databases. When high-quality images were unavailable, only verified reproductions were retained to ensure both reliability and comprehensive coverage of modern and contemporary Korean paintings within the dataset.

Cultural and institutional validation

Following the passing of Lee Kun-hee, the late chairman of the Samsung Group, in 2020, his family made an unprecedented donation of over 23,000 cultural artifacts to the National Museum of Korea (NMK) and the National Museum of Modern and Contemporary Art (MMCA)—an event widely described as “the donation of the century.” Of these, 21,600 works were entrusted to the NMK and 1,488 to the MMCA⁷⁷. This monumental bequest formed the basis for *The Lee Kun-hee Collection: Masterpieces of Korean Art*, in which the eleven artists featured in our dataset were also prominently represented⁷⁸. Their inclusion in this nationally curated exhibition affirms their acknowledged status as canonical figures in modern and contemporary Korean painting, thereby minimizing potential sampling bias and reinforcing both the cultural and academic legitimacy of the dataset.

Beyond establishing cultural and institutional validation, it is equally important to strengthen the linkage between quantitative findings and art-historical interpretation. Rather than merely identifying which computational methods perform effectively for particular artists, analysis should integrate concrete visual evidence—such as close-up examinations of brushstrokes—to elucidate *why* these methods succeed. Such integration

Table 1 | Composition of the dataset and artist-specific representative images












Artist name	Number of images	Representative image
Kim Ki-chang	100	 <p>© NEWSIS.COM All rights reserved.</p>
Kim Whan-ki	100	 <p>© Whanki Foundation-Whanki Museum</p>
Do Sang-bong	100	 <p>© NEWSIS.COM All rights reserved.</p>
Park Soo-keun	100	 <p>© Park Soo Keun Institute</p>
Byun Kwan-sik	100	 <p>© NEWSIS.COM All rights reserved.</p>
Byun Jong-ha	100	 <p>© NEWSIS.COM All rights reserved.</p>
Yoo Young-kuk	100	 <p>© YOO YOUNGKUK ART FOUNDATION. All right Reserved.</p>
Lee Sang-beom	100	 <p>© 2024 Artnet Worldwide Corporation. All rights reserved.</p>
Lee Jung-seob	100	 <p>© 2021 National Museum of Modern and Contemporary Art, Korea</p>
Chang Uc-chin	100	 <p>© 2024 BY CHANGUCCHIN FOUNDATION. ALL RIGHTS RESERVED.</p>
Chun Kyung-ja	100	 <p>© Seoul Metropolitan Government all rights reserved. © NEWSIS.COM All rights reserved.</p>

Table 2 | Artists classification by stylistic orientation

Stylistic Orientation	Artists	Stylistic Features
Realism	Do Sang-bong (1902–1977) Park Soo-keun (1914–1965)	Rooted in traditional realism; depicted everyday life and national sentiment with warmth and authenticity
Ink painting	Lee Sang-beom (1897–1972) Byun Kwan-sik (1899–1976)	Developed distinctive schools of ink landscape painting (Cheongjeon style, Sojeong style), modernizing traditional aesthetics
Abstraction	Kim Whan-ki (1913–1974) Yoo Young-kuk (1916–2002) Byun Jong-ha (1926–2000)	Leaders of postwar abstraction and experimental art; combined Korean motifs, geometric abstraction, and spatial innovation
Expressionism/ Individual Styles	Lee Jung-seob (1916–1956) Chun Kyung-ja (1924–2015) Chang Uc-chin (1917–1990)	Strongly personalized idioms: expressive line and form, vivid colors, lyrical and simplified visual language
Folk-inspired/ Hybrid style	Kim Ki-chang (1914–2001)	Reinterpreted traditional folk painting and genre painting through his unique “Foolish Paintings” series

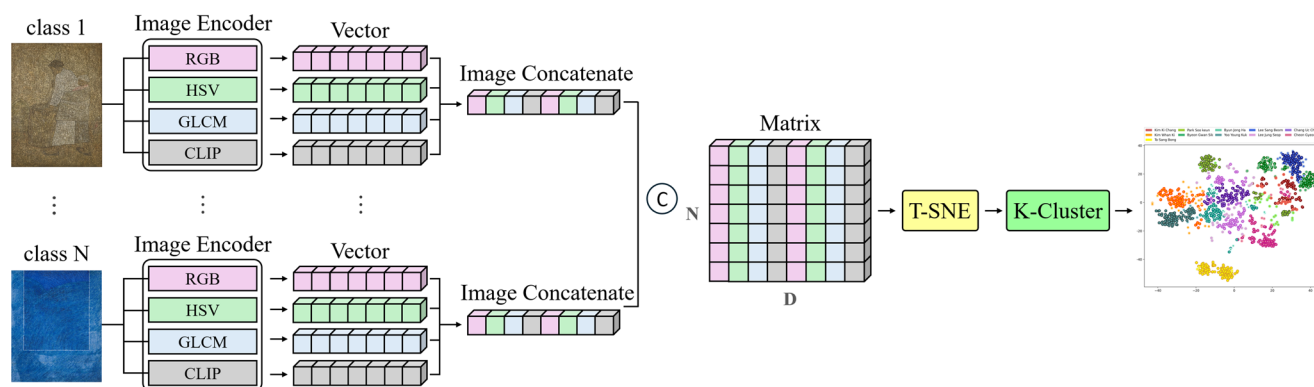


Fig. 1 | Overall architecture for clustering. The proposed framework extracts complementary features (RGB, HSV, GLCM, and CLIP) from each image, concatenates them into a unified representation, and applies dimensionality reduction (t-SNE) followed by K-means clustering to group visually and semantically similar images.

produces richer, more explanatory interpretations that bridge the gap between algorithmic outcomes and aesthetic understanding. This perspective aligns with prior studies such as *The Shape of Art History in the Eyes of the Machine*²¹, which demonstrated that machine-learned stylistic features correspond closely with established art-historical taxonomies, and *Computational Analysis of Content in Fine Art Paintings*⁴⁰, which revealed meaningful co-occurrence patterns among pictorial subjects through quantitative methods.

Stylistic and temporal diversity

The dataset was designed to balance both stylistic and temporal diversity, thereby reducing sampling bias and enhancing representativeness. The selected artists encompass abstraction, realism, traditional ink painting, expressionism, and folk-inspired hybrid practices, as summarized in Table 2. This classification confirms that the corpus spans a broad spectrum of artistic practices reflective of modern and contemporary Korean art. For example, Lee Sang-beom and Byun Kwan-sik are widely recognized for mastery of traditional ink painting, Kim Whan-ki represents postwar abstract composition, and Park Soo-keun exemplifies modern realism. Moreover, the dataset extends across the twentieth century, covering early modernist tendencies as well as later experimental approaches.

All works were digitized at high resolution, and image quality and authenticity were manually verified. Metadata—including artist name, year of production (when available), and artistic category (e.g., sketch, watercolor, oil, ink painting)—were documented to provide contextual information and to enable more granular computational analysis. To ensure reproducibility, we recorded the data-collection workflow and the criteria for artist and work selection. By clarifying the distribution of styles and categories, the dataset mitigates concerns about sampling bias and strengthens the reliability of subsequent analyses.

Proposed framework

This study proposes an analytical framework integrating multiple image feature extraction techniques, dimensionality reduction, and clustering to analyze artworks by modern and contemporary Korean painters (Fig. 1). The framework consists of the following stages.

First, input data comprise images categorized by artist, with object regions cropped according to pre-existing annotations before processing. The collected artworks are then transformed into diverse visual feature vectors using multiple encoders. Four encoding methods—RGB means values⁷⁹, HSV means values⁵¹, color histograms, the Gray-Level Co-occurrence Matrix (GLCM)⁸⁰ for texture analysis, and CLIP embeddings⁸¹—were employed to extract complementary representations. RGB and HSV function as color spaces, GLCM captures texture, and CLIP provides high-level semantic features.

All feature types were L2-normalized and concatenated with equal weight to form a single multimodal feature vector for each image. Each vector was stored with its corresponding filename to ensure traceability and facilitate retrieval of representative or misclustered samples. Equal-weight fusion was adopted for three reasons: (1) the limited dataset size and heterogeneity could cause learned weights or attention mechanisms to overfit; (2) equal weighting maintains interpretability by transparently balancing color, texture, and semantic cues; and (3) it establishes a stable baseline for evaluating the contribution of each modality. Comparative ablation results are presented in Table 4.

Subsequently, all feature vectors were aggregated into a feature matrix representing the dataset’s overall visual characteristics. To visualize the high-dimensional structure, we applied t-SNE for dimensionality reduction, followed by K-means clustering to identify typological similarities and group-level relationships among images. Finally, to assign stylistic labels without task-specific training (e.g., realism, ink painting, abstraction), a

Table 3 | Prompts used for automatic image captioning

Stylistic orientation	prompt
Realism	“a realistic, naturalistic painting with lifelike rendering and accurate light and shadow” “a painting in the style of realism, faithful perspective and fine detail” “자연주의적이고 사실적인 묘사, 정확한 명암과 원근을 갖춘 사실주의 회화”
Ink painting	“an East Asian ink wash painting with calligraphic brushwork on paper” “traditional ink wash painting with graded washes and calligraphic strokes on rice paper” “수묵화: 먹의 번짐과 필획, 종이에 표현된 동양화적 붓질”
Abstraction	“a non-representational abstract painting with geometric forms or color fields” “abstract art: minimal forms, grids, and color fields without recognizable subjects” “추상 회화: 비구상 형태, 기하학적 구성, 색면으로 이루어진 작품”
Expressionism/ Individual Styles	“an expressive painting with vigorous gestural brushstrokes and emotional distortion” “expressionism: impasto texture, dramatic strokes, and distorted figures” “표현주의: 격정적 붓질, 왜곡된 형태, 강한 감정 표현의 회화”
Folk-inspired/ Hybrid style	“a painting with vernacular folk motifs, decorative traditional patterns, and hybrid style” “folk-inspired hybrid style: symbolic animals, decorative motifs, and craft-like patterns” “민화 및 전통 문양에서 영감을 받은 하이브리드 스타일의 회화”

zero-shot vision-language approach was used, matching image and text-prompt embeddings within a shared semantic space via cosine similarity.

CLIP is a vision–language model trained on large-scale image–text pairs that aligns both modalities within a shared embedding space. In our pipeline, we use the pretrained image encoder to extract semantic embeddings only, without any supervised fine-tuning⁸². t-SNE is a nonlinear dimensionality reduction method that preserves local neighborhood structure⁸³. It is used to obtain a low-dimensional representation for visualization and as input for clustering. Because the procedure is unsupervised, labels are attached post hoc by majority mapping within each cluster, and we report clustering accuracy rather than supervised classification accuracy. This capability underpins the zero-shot style assignment described earlier.

Upon image input, feature vectors are generated through four modules—RGB, HSV, GLCM, and CLIP. The RGB module extracts color composition through red, green, and blue channels, while the HSV module computes mean and standard deviation values for each color component, reflecting edge count, dark-pixel ratio, symmetry, and average values in hue, saturation, and value spaces. GLCM, a statistical texture analysis technique, characterizes textures by quantifying the frequency of specific pixel-pair occurrences at predefined spatial relationships, thereby enabling detailed statistical measurement of texture. Lastly, CLIP captures semantic associations between text and images by encoding each image into an embedding vector and linking it with textual descriptions. The model calculates cosine similarity between image and text embeddings, selecting the most relevant description. We build upon this capability to assign style labels through zero-shot matching of image and text-prompt embeddings via cosine similarity, without any additional training. In this study, pretrained CLIP encoders are employed to improve feature representation and ensure consistency across the dataset.

The clustering module employs K-means clustering on low-dimensional feature vectors projected by t-SNE, automatically grouping modern and contemporary Korean paintings by typological similarity. Feature vectors from all four modules are concatenated into a unified vector per image and vertically stacked across the dataset to construct a feature matrix. This high-dimensional matrix is reduced to two dimensions—*x* and *y* axes—through t-SNE, enabling visual representation of feature space. By integrating four complementary modules, the matrix effectively captures both color and texture attributes, aiding clustering and evaluation through majority-label assignment. The resulting two-dimensional vectors, which preserve essential visual features, serve as inputs for the clustering process.

Subsequently, K-means clustering partitions the dataset into 11 × 20 initial clusters, a value of *K* chosen to capture the diversity of styles and evolving techniques among the eleven modern and contemporary Korean painters represented. These initial clusters maximize granularity by identifying fine stylistic variations at early stages. Post-processing refinement

was then applied by analyzing cluster sizes, removing smaller or redundant clusters, and retaining only the upper half to enhance interpretability. Final clustering was performed using the centroids of the retained clusters, followed by reapplying K-means to produce the final grouping.

The clustering results were evaluated by selecting representative images for each cluster and analyzing the distribution of ground-truth labels (painter names) to assess cluster purity and mean accuracy. The dominant ground-truth label within each cluster was assigned as its representative label, quantifying the degree to which the clusters corresponded to individual painters’ stylistic tendencies. The outcomes were visualized using representative images and captions, providing interpretable insights into the model’s ability to capture stylistic groupings.

In addition, clusters were displayed with circle colors denoting ground-truth labels and outline colors representing clustering results, allowing intuitive verification of label alignment. This visualization approach supported qualitative evaluation of stylistic coherence and misclassification patterns. The hierarchical clustering strategy thus balanced the capture of artistic diversity during the initial partitioning stage with interpretability in subsequent refinements. Given that modern and contemporary Korean painters exhibit both distinctive personal styles and intra-artist variations across periods and subjects, this multistage clustering methodology effectively accommodated the dataset’s heterogeneity and stylistic complexity.

For zero-shot style assignment, we constructed label-specific prompt sets in both English and Korean, incorporating descriptors related to medium, brushwork, composition, and degrees of abstraction or expression. For each style label, multiple text prompts were encoded using the CLIP text encoder. The resulting prompt embeddings were L2-normalized and averaged to form a single prototype representation per label. Style prediction was performed by computing cosine similarity between the image embedding and each label prototype, followed by a softmax operation across labels to generate probabilistic style assignments. The complete prompt sets employed for each style category are summarized in Table 3.

Results

Implementation detail

In this study, feature vectors were constructed by extracting image embeddings using the pre-trained CLIP model ViT-B/32 and integrating additional color and texture information. All experiments were performed on an NVIDIA RTX 3090 GPU using Python 3.10 on Ubuntu 22.04 with CUDA 12.2. Input images were normalized to a resolution of 256 × 256 pixels, and a batch size of 32 was applied for CLIP embedding extraction.

Each feature vector consisted of a 512-dimensional CLIP embedding (ViT-B/32) concatenated with RGB and HSV channel means (three each) and 23-dimensional GLCM texture statistics, yielding a 541-dimensional representation per image. All components were z-score normalized prior to clustering. For comparison, additional features were computed—768-

Table 4 | Ablation results for multimodal feature fusion in clustering-based artist identification

Method	Clustering accuracy (%)	Silhouette	Notes
CLIP	82.2	0.4047	Single
CLIP + RGB	82.0	0.4105	Pairwise fusion
CLIP + HSV	81.3	0.4065	Pairwise fusion
CLIP + Histogram	51.0	0.3791	Pairwise fusion
CLIP + LBP	68.8	0.3618	Pairwise fusion
CLIP + GLCM	73.7	0.4040	Pairwise fusion
CLIP + RGB + HSV	81.8	0.4112	Triple fusion
CLIP + LBP + GLCM	73.6	0.3678	Triple fusion
CLIP + RGB + HSV + GLCM (Ours)	82.4	0.4180	Equal-weight fusion

The bold values indicate the best-performing results for each comparison setting, representing the highest clustering or classification accuracy achieved among the evaluated methods or feature combinations.

dimensional color histograms and 529-dimensional Local Binary Pattern (LBP) descriptors—which served as baselines but were excluded from the proposed multimodal fusion. For t-SNE dimensionality reduction, the learning rate was set to 200 and perplexity to 30, controlling the step size of the embedding updates and the effective neighborhood size, respectively.

To assess stability, the full t-SNE and K-means pipeline was executed with three independent random seeds, producing clustering accuracies of 82.3%, 82.4%, and 82.4%. The mean accuracy was 82.37% with a standard deviation of approximately 0.06 percentage points, and the canonical run is reported as 82.4%. K-means clustering was initialized with a number of clusters equal to twenty times the number of artists, followed by filtering to retain the top 50% by cluster size. Clustering accuracy was calculated under post hoc majority-label assignment as the proportion of images whose ground-truth labels matched the dominant label within each cluster. All experiments were conducted with controlled random seeds to ensure reproducibility.

Ablation study

Quantitative performance evaluation followed the method of Fränti and Sieranoja, in which the most frequent label within each cluster was assigned as the representative label, and clustering accuracy was defined as the proportion of images matching that label⁸⁴. This metric effectively quantifies cluster purity and is particularly suitable when ground-truth labels are clearly defined (e.g., clustering by artist or artwork type), enabling direct comparison with prior feature-extraction baselines. Previous studies have primarily relied on RGB, HSV, and histogram-based color features⁸⁵, along with LBP and GLCM for texture. In contrast, our multimodal approach integrates these visual cues with CLIP’s semantic representation, and accompanying silhouette analysis further validates that the discovered clusters are structurally meaningful beyond mere label agreement⁸⁶.

The clustering results are summarized in Table 4, with the best results highlighted in bold. Clustering was conducted based on the visual characteristics of the paintings, and intra-cluster consistency was assessed through both quantitative and visual analyses. Specifically, features extracted from the pre-trained CLIP model were combined with RGB, HSV, histogram-based color features, and texture features derived from LBP and GLCM. These were projected into a low-dimensional space using t-SNE and subsequently clustered into eleven primary groups using the K-means algorithm.

As reported in Table 4, the proposed equal-weight fusion of CLIP, RGB, HSV, and GLCM achieved the highest clustering accuracy of 82.4%, along with the best silhouette score of 0.4180, indicating tighter and better-separated clusters than alternative variants. The CLIP-only baseline achieved 82.2% accuracy with a silhouette score of 0.4047, while the addition

Table 5 | Performance comparison between hierarchical linkages and the proposed K-means method (k = 11)

Method	Clustering accuracy (%)	Silhouette
Hierarchical - ward	54.5	0.0642
Hierarchical - average	10.4	0.0212
Hierarchical - complete	38.9	0.0523
Hierarchical - single	10.1	0.0158
K-means (Ours)	82.4	0.4180

The bold values indicate the best-performing results for each comparison setting, representing the highest clustering or classification accuracy achieved among the evaluated methods or feature combinations.

of color and texture cues provided a more stable and balanced representation. Conversely, inclusion of histogram features decreased accuracy to 51.0% and lowered the silhouette score to 0.3791, suggesting that coarse color distributions introduced noise and disrupted semantic alignment. The CLIP + LBP combination yielded 68.8% accuracy with a silhouette score of 0.3618, demonstrating limited benefit. Overall, silhouette trends closely paralleled accuracy—methods that improved accuracy also enhanced silhouette values—supporting the conclusion that multimodal fusion strengthens both intra-cluster cohesion and inter-cluster separation.

While quantitative metrics are standard in machine learning, qualitative interpretation remains essential in artistic-data analysis. Clustering accuracy provides an objective measure for comparison, and when combined with silhouette scores, reinforces the conclusion that the resulting clusters are organized around distinct, interpretable visual characteristics.

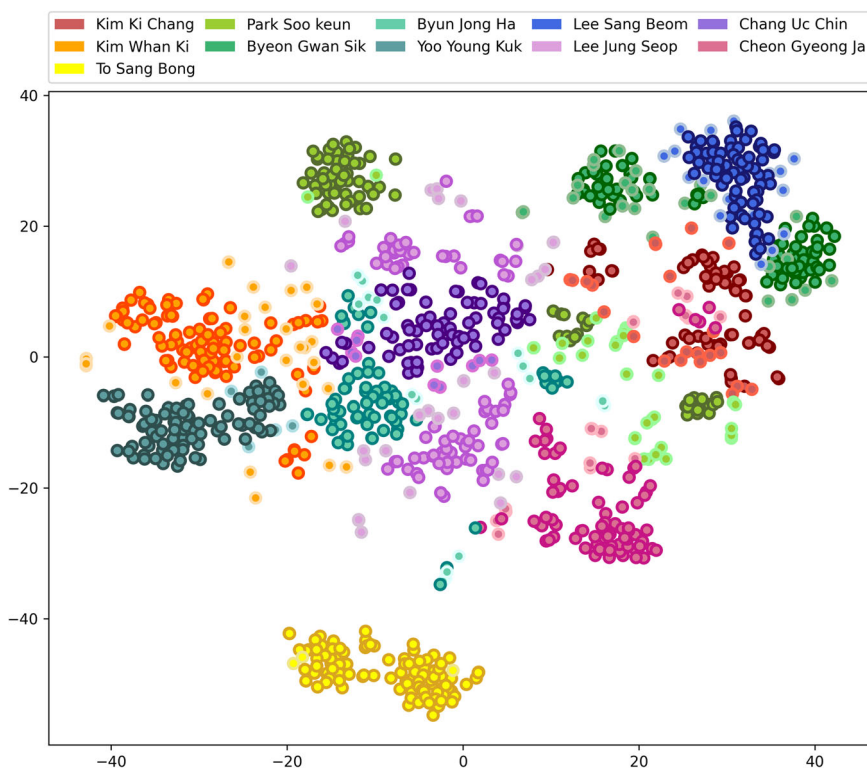
Table 5 compares the performance of different hierarchical linkage methods with the proposed K-means approach (k = 11). Among the hierarchical variants, Ward’s linkage achieved the best accuracy of 54.5% and a silhouette score of 0.0642. However, all hierarchical methods performed substantially below K-means, indicating that centroid-based partitioning is considerably more effective for separating visually diverse artworks. This result suggests that the stylistic variability and high-dimensional feature structure of the dataset are better captured by centroid-oriented clustering than by distance-based hierarchical aggregation.

We further compared traditional color- and texture-based methods with the proposed approach, stratified by artist and artwork type. Color-based models (RGB, HSV) achieved high accuracy for artists whose practices are distinguished by characteristic palettes. For example, Yoo Young-kuk—known for bold primary colors and geometric color planes—reached 85.0% with RGB and 92.0% with HSV. Similarly, Kim Whan-ki recorded 72.0% (RGB) and 75.0% (HSV). Do Sang-bong also attained exceptionally high recognition rates under color-based analysis, reflecting the recurring tonalities and stable harmonies of his still-life works. Taken together, these results indicate that color-centric models are particularly effective when an artist’s signature palette is a primary stylistic determinant, whereas alternative cues (e.g., texture or semantics) may be required when color is less diagnostic.

Conversely, histogram-based methods yielded comparatively low accuracy across all cases, indicating that coarse color-distribution statistics are insufficient to capture the nuanced stylistic and painterly characteristics of individual artists. Notably, Byun Jong-ha and Chang Uc-chin fell below 30%, underscoring the limitations of relying solely on basic color-frequency comparisons for artist identification.

The LBP method—which encodes local texture and tonal transitions—performed strongly for Do Sang-bong, likely reflecting the detailed brushwork and shading typical of his still-life paintings. Yoo Young-kuk also benefited from LBP due to the pronounced contrast at boundaries between color fields. By contrast, GLCM—capturing spatial co-occurrence of gray levels—achieved 100% accuracy for Do Sang-bong, consistent with the stable compositions and recurring textural patterns in his still lifes.

Fig. 2 | Clustering results of Korean modern and contemporary artists' paintings.



In contrast, artists such as Kim Ki-chang and Lee Jung-seob, whose works exhibit more spontaneous, dynamic composition and gesture, showed relatively lower performance under pattern-based analysis. Similarly, the experimental and complex techniques of Byun Jong-ha, as well as the simplified, planar compositions characteristic of Chang Uc-chin, were not optimally captured by any single traditional method, resulting in lower overall accuracy.

Taken together, these results suggest that color-centric techniques are most effective when distinctive palettes are diagnostic, whereas texture- and pattern-based methods better suit artists whose practices foreground compositional structure and surface articulation. This highlights the importance of method selection aligned with the inherent visual characteristics of each artist's oeuvre.

Summary

Conventional approaches achieved relatively high accuracy for certain artists but degraded substantially for those employing distinctive or unconventional expressive techniques. This limitation stems from the sensitivity of traditional features to specific color combinations or local texture patterns, which fail to capture the holistic visual context of paintings. In contrast, the proposed method—integrating CLIP-based semantic features with color and texture cues—delivered stable and consistent accuracy across all artists and artwork categories, mitigating artist- or style-specific bias and overcoming the constraints of conventional feature sets.

Visualization of clustering results

Figures 2 and 3 were revised with clearer cluster labels and explicit highlighting of misassigned samples. The updated visualizations show that most clusters are well separated, while boundary overlaps occur where stylistic similarities exist. A comparative analysis of these boundary cases has been added to strengthen the persuasiveness of the results.

Figure 2 visualizes the K-means grouping into eleven primary clusters. Each artwork is represented as a point; dark-bordered points indicate images correctly assigned under majority-label mapping, whereas light-bordered points indicate misassignments, enabling straightforward

verification of assignment accuracy. The high proportion of dark-bordered points indicates that, despite being an unsupervised method, the system effectively identifies inter-artwork similarity and groups images by artist with high fidelity. This finding underscores the model's capacity to distinguish artists based solely on image information and highlights the value of machine-learning-driven analysis in heritage contexts.

Figure 3 presents a schematic layout of representative images closest to each cluster centroid. These exemplars capture the dominant visual characteristics of their respective clusters, which are organized by stylistic attributes such as color composition, texture patterns, and compositional structure. For example, one cluster is primarily composed of traditional ink paintings with monochromatic tonal variations and distinct brushwork textures, whereas another is dominated by abstract works featuring bold primary colors and geometric forms; still-life paintings form a separate, coherent grouping as well. This visualization confirms that the clustering process captures latent visual attributes, producing interpretable mappings between artists and artworks. Inspection of centroid-proximal exemplars demonstrates that the proposed analytical pipeline effectively aggregates visually similar works and offers a transparent lens on stylistic relationships, while the documented boundary overlaps provide actionable targets for qualitative review and art-historical interpretation.

Cluster prototypes and caption-based interpretation

For a more systematic group-level assessment, cluster centroids were examined Table 6 to analyze inter-group differences. These centroids represent the average position of all data points within each cluster, effectively summarizing the dominant color tones and texture characteristics of each group. Table 7 presents the centroids of correctly assigned clusters, showing that the generated captions accurately reflect each group's visual attributes. This finding suggests that both the clustering algorithm and caption-generation model exhibit high reliability in capturing distinctive stylistic features.

Misassignment analysis

As shown in Table 8, certain clusters produced incorrect artist assignments and inaccurate captions, revealing areas for further refinement. Earlier

Fig. 3 | Schematic representation of representative images nearest to each cluster.

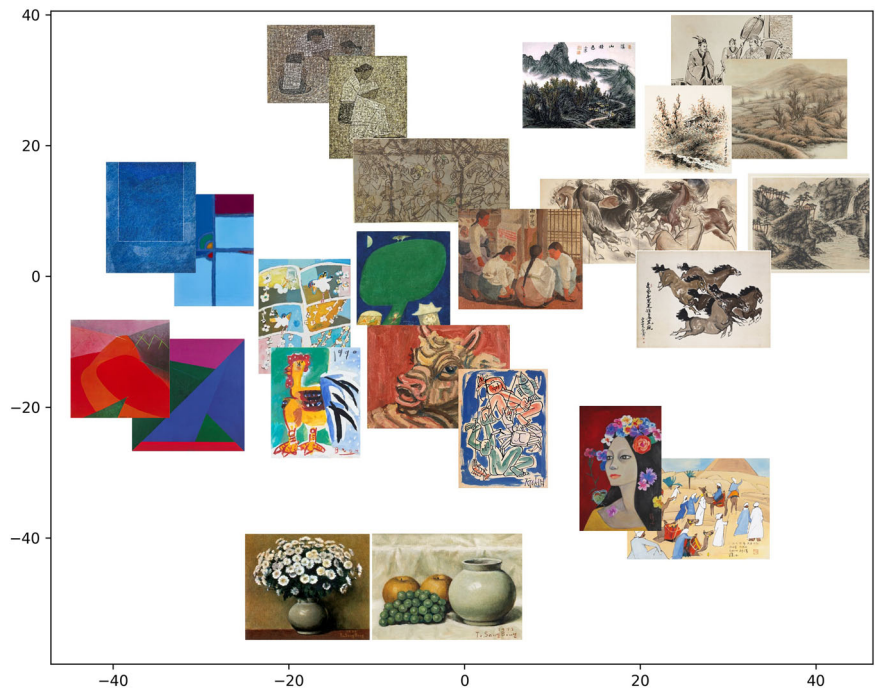


Table 6 | Comparison of model accuracy by artist (unit: %)

Dataset	Method					
	RGB	HSV	Histogram	LBP	GLCM	Ours
Kim Ki Chang	41.0	44.0	33.0	34.0	42.0	73.5
Kim Whan Ki	72.0	75.0	41.0	66.0	76.0	74.8
To Sang Bong	98.0	98.0	37.0	98.0	100	97.0
Park Soo Keun	56.0	56.0	38.0	50.0	56.0	69.1
Byeon Gwan Sik	50.0	49.0	30.0	53.0	51.0	76.0
Byun Jong Ha	49.0	58.0	24.0	36.0	51.0	78.3
Yoo Young Kuk	85.0	92.0	56.0	70.0	93.0	94.9
Lee Sang Beom	65.0	64.0	32.0	63.0	60.0	78.4
Lee Jung Seop	43.0	42.0	27.0	26.0	40.0	74.4
Chang Uc Chin	49.0	47.0	24.0	36.0	60.0	86.4
Cheon Gyeong Ja	67.7	75.8	18.2	65.7	74.8	82.3

All methods use as the backbone; labels denote additional features concatenated to the CLIP embedding. The bold values indicate the best-performing results for each comparison setting, representing the highest clustering or classification accuracy achieved among the evaluated methods or feature combinations.

versions of this paper only noted the occurrence of misassignments; in this revision, we provide a detailed per-cluster breakdown in Table 8. The model most frequently confused artists whose works share overlapping stylistic characteristics. For instance, Kim Ki-chang was occasionally misassigned to Lee Jung-seob, likely due to similarities in brushstroke fluidity and color palette. Likewise, abstract compositions by Kim Whan-ki and Yoo Young-kuk were often grouped together, reflecting shared compositional structures and chromatic strategies. Importantly, such misassignments are not mere errors but indicate meaningful visual commonalities, providing insight into aesthetic proximities between artists. Thus, the computational model not only detects stylistic boundaries but also illuminates inter-artist affinities embedded in the visual corpus.







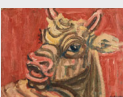
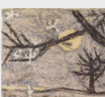

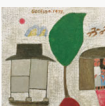
Linking quantitative results with Art-historical interpretation

To strengthen the connection between quantitative outcomes and art-historical interpretation, this study moves beyond simply identifying which

Table 7 | Results showing alignment between cluster centroid and correct answer

Correct answer Image	Cluster centroid Class	Image	Alignment		Remark
			Predict class	Right class	
	To Sang-bong		To Sang-bong	To Sang-bong	Realism
	To Sang-bong		To Sang-bong	To Sang-bong	Realism
	Chang Uc-chin		Chang Uc-chin	Chang Uc-chin	Expressionism/ Individual Styles
	Kim Whan-ki		Kim Whan-ki	Kim Whan-ki	Abstraction
	Kim Whan-ki		Kim Whan-ki	Kim Whan-ki	Abstraction
	Cheon Gyeong-ja		Cheon Gyeong-ja	Cheon Gyeong-ja	Expressionism/ Individual Styles

Table 8 | Results showing alignment between cluster centroid and incorrect answer

Incorrect answer		Cluster centroid		Remarks	
Image	Class	Image	Predict class	Right class	
	Park Soo-keun		Kim Ki-chang	Park Soo-keun	Folk-inspired/Hybrid style
	Lee Sang-beom		Byeon Gwan-sik	Lee Sang-beom	Ink painting
	Lee Jung-seop		Chang Uc-chin	Lee Jung-seop	Expressionism/Individual Styles.
	Lee Jung-seop		Byun Jong-ha	Lee Jung-seop	Expressionism/Individual Styles
	Chang Uc-chin		Kim Ki-chang	Chang Uc-chin	Folk-inspired/Hybrid style

computational methods were most effective. It incorporates concrete visual evidence—such as close-up analysis of brushstrokes and compositional framing—to explain why specific methods succeeded or failed in capturing an artist’s distinctive style. This integrative approach enriches the interpretive dimension of the research, demonstrating that computational analysis can meaningfully align with established art-historical methodologies. By grounding algorithmic results in verifiable visual features, the study ensures interpretive reproducibility and methodological transparency.

Furthermore, this framework illustrates how quantitative and qualitative approaches can function complementarily — revealing deeper insights into the visual language of modern and contemporary Korean painting. It also offers a transferable model for future research in digital art history and heritage science, where data-driven analysis must coexist with the interpretive nuance of visual scholarship.

Discussion

This study proposed a multimodal machine-learning framework for the computational analysis of twentieth-century Korean paintings. By integrating color (RGB, HSV), texture (GLCM), and semantic (CLIP) features and applying unsupervised clustering—with majority ground-truth labels assigned only at evaluation time—the fusion approach achieved an artist-level clustering accuracy of 82.4%, outperforming unimodal baselines. These results demonstrate the utility of machine learning as an analytical instrument for Korean modern and contemporary art, a domain that remains underexplored within digital art history.

Our contributions are twofold. First, we curated a dedicated dataset of 1100 works by 11 representative artists, providing a reproducible foundation for quantitative studies in Korean art. Second, we validated a fusion strategy that combines color, texture, and semantic channels; ablation experiments clarified the contribution of each modality, and misassignment analyses—interpreted within art-historical context—yielded insights into inter-artist stylistic proximities beyond numeric performance alone.

Looking ahead, we envision extending this work by expanding artist and work coverage with richer metadata, incorporating temporal features to track stylistic evolution, and leveraging advanced deep-learning methods (e.g., attention-based fusion and large-scale vision–language models) to enhance recognition and interpretation. Cross-cultural comparisons with Western art datasets will further illuminate shared structures and tradition-specific differences. Taken together, these directions will advance the quantitative study of Korean painting and contribute to computational art history more broadly. By fostering sustained collaboration between art historians and AI researchers, this work lays groundwork for a reproducible, data-driven methodology that complements traditional art-historical interpretation and supports the continued development of digital humanities and heritage science.

Despite the promising findings, this study has several limitations. First, the dataset remains relatively small—comprising 1,100 works from 11 artists—which may constrain the generalizability of the results. Second, clustering stability poses a challenge for artists with fewer representative works or highly diverse stylistic practices. Third, the current framework struggles to capture complex hybrid styles in which traditional ink-wash elements coexist with abstract or modernist techniques. Finally, image captioning was implemented as an exploratory experiment; however, it exhibited a notable gap from the visuo-perceptual analysis and was limited in representing the subjective and symbolic dimensions of the artworks. While the results are meaningful from an engineering standpoint, assigning interpretive significance to automatically generated captions within an art-historical framework remains constrained. Future research should refine the methodology so that visuo-perceptual analyses—grounded in art-historical perspectives—directly yield interpretively valid results.

Data availability

The dataset used in this study consists of paintings by eleven modern and contemporary Korean artists. The image data were provided by multiple institutions, including the Museum of Modern and Contemporary Art (MMCA), the Jang Uc-Chin Foundation, the Seoul Metropolitan Government, the Park Soo Keun Institute, and the Yoo Youngkuk Art Foundation. All copyrights for the images remain with the respective institutions. Due to copyright restrictions, the authors are not permitted to redistribute the raw image files. Researchers who wish to access the dataset may request permission directly from the rights holders through each institution’s official application portal or curatorial/research office.

Received: 7 August 2025; Accepted: 6 January 2026;
Published online: 31 January 2026

References

1. Viswanathan, N. & Stanford. Artist identification with convolutional neural networks. <https://api.semanticscholar.org/CorpusID:198974841> (2017).
2. Kim, M. & Kim, J. Complementary quantitative approach to unsolved issues in art history: Similarity of visual features in the paintings of Vermeer and his probable mentors. *Leonardo* **52**, 164–174 (2019).
3. Shamir, L. et al. Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. *ACM Trans. Appl. Percept.* **7**, 1–17 (2010).
4. Cetinic, E., Lipić, T. & Grgic, S. Fine-tuning convolutional neural networks for fine art classification. *Expert Syst. Appl.* **114**, 107–118 (2018).
5. Saleh, B., Abe, K., Arora, R. & Elgammal, A. Toward automated discovery of artistic influence. *Multimed. Tools Appl.* **75**, 3565–3591 (2016).
6. Bar, Y., Levy, N. & Wolf, L. Classification of artistic styles using binarized features derived from a deep neural network. In *Proc. ECCV Workshops*, 71–84 (2014).
7. Zujovic, J. et al. Classifying paintings by artistic genre: An analysis of features & classifiers. In *Proc. IEEE international workshop on multimedia signal processing* (2009).

8. Saleh, B. & Elgammal, A. A unified framework for painting classification. In *Proc. IEEE International Conference on Data Mining Workshop* (2015).
9. Gultepe, E., Conturo, T. E. & Makrehchi, M. Predicting and grouping digitized paintings by style using unsupervised feature learning. *J. cultural Herit.* **31**, 13–23 (2018).
10. Chen, J. & Deng, A. Comparison of machine learning techniques for artist identification. *Stanford193CS231N Report*; Art Computer Science: Stanford, CA, USA (2018).
11. Berezhnoy, I. E., Postma, E. O. & van den Herik, H. J. Authentic: computerized brushstroke analysis. In *Proc. IEEE International Conference on Multimedia and Expo* (2005).
12. van den Herik, H. J. & Postma, E. O. Discovering the visual signature of painters. In *Future Directions for Intelligent Systems and Information Sciences: The Future of Speech and Image Technologies*. Brain Computers, WWW, and Bioinformatics 129–147 (2000).
13. Falomir, Z., Lledó, M., Sanz, I. & Gonzalez-Abril, L. Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn). *Expert Syst. Appl.* **97**, 83–94 (2018).
14. Lamberti, F., Sanna, A., & Paravati, G. Computer-assisted analysis of painting brushstrokes: digital image processing for unsupervised extraction of visible features from van Gogh's works. *J. Image Video Proc.* **2014**, 53 (2014).
15. Johnson, R. C. et al. Image processing for artist identification. *IEEE Signal Process. Mag.* **25**, 37–48 (2008).
16. van der Maaten, L. & Postma, E. Identifying the real van gogh with brushstroke textons. *White Paper, Tilburg University* (2009).
17. Condorovici, R. G., Vertan, C. & Florea, L. Artistic genre classification for digitized painting collections. *UPB Sci. Bull.* **75**, 75–86 (2013).
18. Stork, D., Coddington, J. & Bentkowska-Kafel, A. Computer vision and image analysis of art. In *Proceedings of SPIE* **7531** (2010).
19. Deng, J. et al. Imagenet: A large-scale hierarchical image database. In *Proc. IEEE conference on computer vision and pattern recognition* (2009).
20. Sigaki, H. Y. D., Perc, M. & Haroldo, R. V. History of art paintings through the lens of entropy and complexity. *Natl. Acad. Sci.* **115**, E8585–E8594 (2018).
21. Elgammal, A. et al. The shape of art history in the eyes of the machine. *Proc. AAAI Conf. Artif. Intell.* **32**, 2183–2191 (2018).
22. Nemade, R., Nitsure, A., Hirve, P. & Mane, S. B. Detection of forgery in art paintings using machine learning. *Int. J. Innov. Res. Sci. Eng. Technol.* **6**, 8681–8692 (2017).
23. Stork, D. Computer image analysis of paintings and drawings: an introduction to the literature. In *Proc. Image Processing for Artist Identification Workshop* (2008).
24. Cetinic, E., Tomislav, L. & Grgic, S. A deep learning perspective on beauty, sentiment, and remembrance of art. *IEEE Access* **7**, 73694–73710 (2019).
25. Ellis, M. H. & Johnson, C. R. Jr. Computational connoisseurship: enhanced examination using automated image analysis. *Vis. Resour.* **35**, 125–140 (2019).
26. Yu, Q. & Shi, C. An image classification approach for painting using improved convolutional neural algorithm. *Soft Comput.* **28**, 847–873 (2024).
27. Mondal, K. & Anita, H. Categorization of artwork images based on painters using CNN. *J. Phys.: Conf. Ser.* **1818**, 012223 (2021).
28. Cervantes, F. et al. Evaluation of CNN models with transfer learning in art media classification in terms of accuracy and class relationship. *Computación y. Sist.* **28**, 233–244 (2024).
29. Zhang, X. & Ding, T. Style classification of media painting images by integrating ResNet and attention mechanism. *Heliyon* **10**, e27178 (2024).
30. Li, Q., Yue, L. & Feng, Q. Research on automatic classification method of artistic styles based on attention mechanism convolutional neural network. *Int. J. Low.-Carbon Technol.* **19**, 2018–2023 (2024).
31. Zeng, L. Improved painting image style classification of ResNet based on attention mechanism. In *Proc. International Conference on Advances in Electrical Engineering and Computer Applications* (2023).
32. Polak, A. et al. Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. *J. cultural Herit.* **26**, 1–11 (2017).
33. Sahu, T., Tyagi, A., Kumar, S. & Mittal, A. Classification and aesthetic evaluation of paintings and artworks. In *Proc. 13th International Conference on Signal-Image Technology & Internet-Based Systems* (2017).
34. Castellano, G. & Vessio, G. A deep learning approach to clustering visual arts. *Int. J. Comput. Vis.* **130**, 2590–2605 (2022).
35. Abas, F. S. & Martinez, K. Classification of painting cracks for content-based analysis. *Mach. Vis. Appl. Ind. Inspection XI* **5011**, 149–160 (2003).
36. Andronache, I. Can complexity measures with hierarchical cluster analysis identify overpainted artwork? *Zenodo* **10**, 1–58 (2023).
37. Baek, S. et al. Identification of Korean Neo-Realism artists through CLIP-based analysis. *J. Korea Soc. Comput. Inf.* **29**, 317–328 (2024).
38. Elgammal, A., & Saleh, B. Quantifying creativity in art networks. In *Proc. 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2015).
39. Saleh, B. & Elgammal, A. Large-scale classification of fine-art paintings: Learning the right metric on the right feature. *Int. J. Digital Art. Hist.* **2**, 70–93 (2016).
40. Kim, E., Xu, K., Elgammal, A., & Mazzone, M. Computational Analysis of Content in Fine Art Paintings. In *Proc. the AAAI Conference on Artificial Intelligence (AAAI-19)* (2019).
41. Agarwal, S., Karnick, H., Pant, N. & Patel, U. Genre and style based painting classification. In *Proc. IEEE Winter Conference on Applications of Computer Vision* (2015).
42. Mensink, T. & Van Gemert, J. The rijksmuseum challenge: Museum-centered visual recognition. In *Proc. International conference on multimedia retrieval* (2014).
43. Kim, D., Xu, J., Elgammal, A. & Mazzone, M. Computational analysis of content in fine art paintings. In *Proc. ICCV* (2019).
44. Li, J., Yao, L., Hendriks, E. & Wang, J. Z. Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**, 1159–1176 (2011).
45. Park, J. W. Secrets of balanced composition as seen through a painter's window: visual analyses of paintings based on subset barycenter patterns. In *ACM SIGGRAPH 2019 Art Gallery* 1–10 (2019).
46. Carneiro, G. Graph-based methods for the automatic annotation and retrieval of art prints. In *Proc. 1st ACM International conference on multimedia retrieval* (2011).
47. Dasgupta, S. & Hossain, M. A. Pattern recognition through brushstroke analysis: an inspection to classify van Gogh from others. *BRAC University*, (2017).
48. Xu, K. Automated artist identification using deep learning and transfer learning techniques. *Sci. Technol. Eng.* **1**, 8 (2024).
49. Zheng, S. Deep learning-based artistic style transformation algorithm in visual communication. *Int. J. Syst. Assurance Eng. Manag.* 1–10 (2024).
50. Wei, N. Research on the algorithm of painting image style feature extraction based on intelligent vision. *Futur. Gener. Comput. Syst.* **123**, 196–200 (2021).
51. Nunez-Garcia, I. et al. Classification of paintings by artistic style using color and texture features. *Comput. Sist.* **26**, 1503–1514 (2022).
52. Hu, B. & Yang, Y. Construction of a painting image classification model based on AI stroke feature extraction. *J. Intell. Syst.* **33**, 20240042 (2024).
53. Wang, Q. Painting style recognition and classification based on convolutional neural network and multi-kernel learning. In *Intelligent Computing Technology and Automation* 1059–1066 (2024).

54. WikiArt. WikiArt Visual Art Encyclopedia. n.d. Web. 3 July 2025. <http://www.wikiart.org>.
55. Ivanova, K. et al. Features for art painting classification based on vector quantization of mpeg-7 descriptors. In *Proc. Data Engineering and Management: Second International Conference* 178–189 (2012).
56. Crowley, E. J. & Zisserman, A. In Search of Art. In Agapito, L., Bronstein, M. M. & Rother, C. (eds.) *Computer Vision - ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part I* 54–70 (2015).
57. Shamir, L. Computer analysis reveals similarities between the artistic styles of Van Gogh and Pollock. *Leonardo* **45**, 149–154 (2012).
58. Karayev, S. et al. Recognizing image style. *arXiv preprint arXiv:1311.3715* (2013).
59. Park, J. S., Kim, S. Y., Yoon, Y. C. & Kim, S. K. Optimizing CNN structure to improve accuracy of artwork artist classification. *J. Korea Soc. Comput. Inf.* **28**, 9–15 (2023).
60. Li, J. & Wang, J. Z. Studying digital imagery of ancient paintings by mixtures of stochastic models. *IEEE Trans. Image Process.* **13**, 340–353 (2004).
61. Jiang, S., Huang, Q., Ye, Q. & Gao, W. An effective method to detect and categorize digitized traditional Chinese paintings. *Pattern Recognit. Lett.* **27**, 734–746 (2006).
62. Lu, G. et al. Content-based identifying and classifying traditional Chinese painting images. In *Proc. Congress on Image and Signal Processing* (2008).
63. Baldrati, A., Bertini, M., Uricchio, T. & Del Bimbo, A. Exploiting CLIP-based multi-modal approach for artwork classification and retrieval. In *International Conference Florence Heri-Tech: The Future of Heritage Science and Technologies*. 140–149, (Springer International Publishing, Cham, 2022).
64. Zhong, S., Huang, X. & Xiao, Z. Fine-art painting classification via two-channel dual path networks. *Int. J. Mach. Learn. Cybern.* **11**, 137–152 (2020).
65. Kim, D., Elgammal, A. & Mazzone, M. Formal analysis of art: proxy learning of visual concepts from style through language models. *arXiv preprint arXiv:2201.01819* (2022).
66. National Museum of Modern and Contemporary Art, Korea. The Most Honest Confession: Chang Ucchin Retrospective. n.d. Web. 14 Sept. 2023. <https://www.mmca.go.kr/exhibitions/exhibitionsDetail.do?menuId=1030000000&exhid=202302150001627>.
67. Joo, S. A Study on the Sublime in YOO YOUNGKUK's Abstract Paintings: Focusing on works since the 1960s. *Department of Aesthetics & Art History, Graduate School of Chosun University, Gwangju, Korea* (2023).
68. Kim, H. S. Study on the Colors of Kim Whan-ki's Painting. *J. Art. Theory Pract.* **3**, 155–172 (2005).
69. Kim, M. J. About Originality on Materials and Techniques in Lee Jung Seob's Painting. *J. Korean Mod. Contemp. Art. Hist.* **32**, 83–117 (2016).
70. Eom, S. M. A Unique Artistic World: Park Soo-keun's Concave and Convex Technique. *One Art World* 30–37 (2022).
71. Park, J. M. The concept of abstraction in New Realism group: New interpretation on geometric abstraction from its own point of view. *Korean Bull. Art. Hist.* **38**, 240–276 (2012).
72. Choi, B. The changing characteristics and meaning of Woonbo Kim, Ki-Chang's 'The Foolish Painting Style. *Orient. Art.* **33**, 233–256 (2016).
73. Lee, M. S. A Study on the Cheon, Kyung-Ja's art works. *Graduate School of Jeju University, Jeju, Korea*. (2011).
74. Lee, Y. U. A study on Byeon kwan sik's Mt. Geumgang paintings. *Art. Hist. Cult. Herit.* **1**, 9–34 (2012).
75. National Museum of Modern and Contemporary Art, Korea. "Today, This Work: Lee Sangbeom | Early Winter | 1926." n.d. Web. 1 June 2023. <https://www.mmca.go.kr/digitals/digitalMovInfo.do?mbld=202306010001268>.
76. Hur, N. Y. The Sensibility Narrative through figuration by Byun Jong-Ha. *Assoc. Biogr. Art. Hist. Korea* **5**, 265–287 (2009).
77. Rhee, B. A., Seo, Y. E., Ro, Y. U., & Kim, G. H. A Study on Visitor Engagement of the Audio Guide and Curating-bot at Museum of Modern and Contemporary Art. In *Proc. the 2024 Summer Conference of the Korean Society of Computer and Information* (2024).
78. Park, M. Lee Kun-hee Collection Masterpieces of Korean arts: Its significance and analysis of key works. *MMCA Stud.: Exhibition Histories* **2021**, 192–211 (2021).
79. Maxwell, J. C. Experiments on colour as perceived by the eye, with remarks on colour-blindness. *R. Soc. Edinb.* **3**, 299–301 (1857).
80. Siqueira, D., Roberti, F., Schwartz, W. R. & Pedrini, H. Multi-scale gray level co-occurrence matrices for texture description. *Neurocomputing* **120**, 336–345 (2013).
81. Nukrai, D., Mokady, R. & Globerson, A. Text-only training for image captioning using noise-injected clip. *arXiv preprint arXiv:2211.00575* (2022).
82. Radford, A. et al. Learning transferable visual models from natural language supervision. *arXiv preprint arXiv:2103.00020* (2021).
83. Maaten, L. V. D. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008). **Nov.**
84. Fränti, P. & Sieranoja, S. Clustering accuracy. *Appl. Comput. Intell.* **4**, 24–44 (2024).
85. Özlü, Ahmet. Color Recognition. GitHub, n.d. Web. 3 July 2024. https://github.com/ahmetozlu/color_recognition.
86. Shahapure, K. R., & Nicholas, C. Cluster quality analysis using silhouette score. In *Proc. 2020 IEEE 7th international conference on data science and advanced analytics (DSAA)* (2020).

Acknowledgements

This research was supported by Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2023 (Project Name: Acquisition of 3D precise information of microstructure and development of authoring technology for ultra-high precision cultural restoration, Project Number: RS-2023-00227749, Contribution Rate: 90%) and by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (RS-2021-II211341, Artificial Intelligence Graduate School Program, Chung-Ang University, Contribution Rate: 10%).

Author contributions

S.H. took primary responsibility for drafting the manuscript, designing the experiments, and conducting the main model experiments. S.J., S.E., and Y.M. contributed to data collection, data refinement, experimental interpretation, and the review of prior research. J.W. was responsible for the experimental design and interpretation of the results. B.A. supported the data refinement, validation of experimental results, and overall quality assurance. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Bo-A Rhee.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026