Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
Special Issue on Art, Design, and Media Convergence in the Digital Era: An Analytical Study of Creative Practices 7(7s)

FINE-GRAINED VISUAL ART ANALYSIS VIA YOLOV9 AND K-MEANS

Huafei Ma [1], Tatiya Theppituck [2]

[1] Faculty of Architecture Art and Design, Naresuan University, Phitsanulok, Thailand
[2] Faculty of Architecture Art and Design, Naresuan University, Phitsanulok, Thailand

Received 25 January 2026
Accepted 22 March 2026
Published 19 May 2026

Corresponding Author
Huafei Ma,
huafei.ma@outlook.com

DOI
10.29121/shodhkosh.v7.i7s.2026.8199

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

Using the most recent YOLOv9 model and optimized K-means clustering for adaptive feature extraction and objective classification, this research suggests a universal framework for fine-grained visual art analysis. With an emphasis on eliminating subjectivity in conventional visual art analysis, the framework converts qualitative artistic qualities into quantifiable parameters by quantifying 12 fundamental morphological features, such as line curvature (σ=0.85), symmetry index (using Hu moments), and contour regularity (using Fourier descriptors). Tested on a variety of 1,000 high-resolution images (including jade carvings, ceramic patterns, and pieces of murals), the framework partitions samples into five different categories: Normative Geometric, Free-Form Geometric, Natural Bionic, Minimalist Line, and Symbolic Pattern. It achieves automated clustering with an overall accuracy of 92.7% and a silhouette coefficient of 0.91. Three art historians' cross-validation shows 94.8% consistency with expert classifications, greatly exceeding conventional AI techniques (e.g., ResNet50 with 84.2% accuracy on the same dataset). In support of applications in digital archiving, stylistic evolution research, and cultural heritage preservation, this work develops a strong technical instrument for extensive digital analysis of visual arts.

Keywords: Fine, Grained Visual Art, YOLOv9, Feature Quantification, K, Means Clustering, Cultural Heritage Technology

1. INTRODUCTION

Traditional carvings, pottery patterns, paintings, and artistic motifs are examples of fine visual arts that are essential for preserving cultural heritage. The historical tales and aesthetic values of various civilizations are reflected in them. One major issue that all of these art forms have faced is that, for a very long time, their study has depended on subjective human explanations. Qualitative descriptors of characteristics like "symmetry," "line smoothness," or "motif density" have become the most common in scholarly discourse. Due to the inherent limits presented by this reliance on human knowledge, rigorous large-scale research has been hampered.

First of all, discrepancies have been introduced by subjectivity. The absence of standardized measurement criteria for artistic qualities has caused classification discrepancies among specialists. For example, a researcher would call a pottery dish "naturalistic" if it has such delicate and asymmetrical floral motifs. Cross-disciplinary study comparisons

are made more difficult by the fact that another researcher would label it as "semi-abstract" [1,2]. According to a recent poll of 15 cultural heritage academics, only 68.3% of the scholars agreed on the styles of 50 sample pieces of art. This demonstrates unequivocally how vital it is to create an impartial framework [3].

Second, for huge datasets, manual analysis is not particularly practical. For instance, a thorough visual examination of a single cultural artifact, such as a ceramic fragment, a mural fragment, or a jade sculpture, could take several hours. When there are more than a few hundred artifacts in the dataset, manual inspection is no longer feasible. Due to this impassable barrier, research has been limited to extremely tiny and dispersed samples, making it more difficult to identify more general stylistic tendencies throughout time periods or geographical areas [4,5].

The advancement of artificial intelligence (AI) and computer vision has presented the prospect of change in the current environment. With the aid of advancements in deep learning, sophisticated models such as YOLOv9 are especially skilled at identifying and measuring extremely specific visual elements, such as the intricate textures or the geometric forms. All of this is possible with unprecedented stability and accuracy [6,7]. According to the Reference, these technologies have started to automate tasks like pottery pattern identification [8] and mural style grouping [9] in the field of cultural heritage study. This method permits extensive analysis work and lessens the impact of subjective elements. But there are still some extremely important problems today: Instead of using adaptive feature engineering or fuzzy perception clustering, the current approaches either rely on extremely basic image processing techniques, like edge detection, or solely concentrate on very restricted, domain-specific applications, like evaluating only one type of workpiece [10,11]. The issue of cross-regional distribution offset and label scarcity are the two primary drawbacks of creative images when compared to natural photos [12].. The structural semantics of the photos will be overlooked because of an over-reliance on the texture style of the photos as the basis for judgment, according to pertinent literature. This will also result in phenomena like limited generalization across data sets if pre-tests are conducted using photos or some recognition models are directly transferred to related style fields like painting [13,14]. These days, a universal framework that may use cross-disciplinary and quantitative aspects to examine different visual artworks is especially needed. It ought to connect the meticulousness of humanistic analysis with the rigor of artificial intelligence technology.

There are several forms of jade carving, including geometric, natural, and symbolic styles. It offers a very compelling test case for this architecture because of its intricate and sensitive nuances. According to their study, jade carving is a perfect place to start when confirming universal approaches since it encapsulates the larger difficulties that fine-grained visual art research faces [15,16]. The research presented in this paper overcomes the drawbacks of conventional art analysis, which only offers qualitative descriptions, and creates a 12-dimensional quantifiable feature system that can be used to a variety of visual art forms, such as murals, ceramics, jade carving, and more. This research then uses this framework to automatically extract features in the four core dimensions:

1) Contour regularity, which measures how closely a form resembles a conventional geometric shape using the Fourier descriptor.
2) Symmetry index: this study computes the bilateral or radial symmetry state using the Hu moment, which goes from 0 (complete asymmetry) to 1 (complete symmetry) [3];
3) Line curvature: Gaussian kernel analysis is used in this study to assess the carved lines' fluidity or stiffness. Here, σ is assumed to be 0.85, which is the normal smoothness of the simplified jade;
4) In terms of texture complexity, this work uses the entropy of the gray-level co-occurrence matrix to differentiate between the ordinary surface and the fine pattern surface. A symbolic pattern is indicated when the entropy is higher than 3.2 [17].

Objective comparisons of works, dynasties, and geographical locations are made possible by this transition from qualitative description of objects to quantitative measurement. We can use statistical techniques to confirm if the Song Dynasty's simplified jade carvings have a lower symmetry index than the Han Dynasty's.

Long-standing expert subjective assessment has been the basis for visual art analysis, which has resulted in differences in how the same piece of art is classified (e.g., the distinction between 'naturalistic' and 'abstract'). This study quantifies qualitative art qualities into 12 factors in an attempt to remove this subjective bias using a framework based on the YOLOv9 model and the K-means clustering method. Each artwork may be consistently and reliably analyzed using the framework's quantifiable properties, which include symmetry, line curvature, and texture complexity. These measured factors maintain high accuracy on large-scale datasets and guarantee consistency in classification among many experts.

The suggested 12-dimensional feature system is conceptually developed from fundamental visual-art dimensions, such as formal order, symmetry, line expressiveness, and symbolic density, rather than treating computational features as exclusively data-driven creations. Our approach provides quantifiable bounds that can be statistically compared across times and themes, and it lessens the dependence on solely qualitative experiences by providing a quantitative 12-dimensional feature system [4].

Strong agreement between the suggested framework and human evaluations is shown by cross-validation against expert annotations, indicating that it has the potential to provide consistent stylistic analysis across sizable datasets. The traditional manual process is inferior to this high level of uniformity. Additionally, it can manage huge datasets with relative ease. In ten minutes, it can categorize 1,000 photos. A group of three researchers would need roughly forty hours to finish this assignment if done by hand.

By combining computer vision methods with art-historical analytical standards, this work takes an overtly multidisciplinary approach.

From a theoretical perspective, this study adds rigor to the Jade carving research methodology. In order to examine how the style changes using more objective criteria, it leverages the previously more subjective qualities to be represented in numbers. The suggested methodology enables statistically supported comparisons of style variation across time periods and geographical areas by converting qualitative aesthetic notions into repeatable numerical descriptors. Furthermore, the five-cluster classification approach provides a uniform framework since it is based on statistical patterns rather than human judgment. This paradigm makes it possible to do multidisciplinary research, such as examining the connections between stylistic shifts and socioeconomic variables like imperial sponsorship or trade, as well as cross-cultural analysis, such as contrasting Japanese minimalism netsuke carving with Chinese simplified jade carving.

The suggested methodology offers a scalable instrument for jade artifact stylistic analysis and digital documentation. Whether it is a private collection or a museum, the YOLOv9 model's powerful function can extract the features of jade carvings to swiftly locate and identify them and create a digital archive. While maintaining the crucial role of expert judgment, the framework is intended to serve as a decision-support tool for scholars and museums, enabling digital archiving and exploratory analysis.

The fact that this research serves as a solid reference for the use of artificial intelligence in other conventional art domains is actually one of its most significant functions. Traditional arts like porcelain and calligraphy, for instance, can be studied in a similar way. Overall, the study shows that by enhancing analytical consistency and scalability, artificial intelligence can complement humanistic inquiry rather than replace it.

2. METHODOLOGY

Bilateral and radial symmetry indices are used to quantify classical aesthetic concepts of harmony and balance, allowing for an objective comparison of expressive and ritualized forms. Particularly in forms that emphasize simplicity or expressive flow (such as the "Han Badao" tradition), line-related characteristics like curvature, curvature variance, and continuity are intended to represent carving techniques and artistic intentions. Last but not least, motif density and texture complexity convert narrative richness and symbolic ornamentation–long stressed in art-historical interpretation–into quantifiable visual characteristics. Visual art analysis concepts directly influence feature selection and interpretation through this theory-driven mapping, guaranteeing that the resulting clustering results are both statistically sound and have art-historical significance.

1) Verified heritage-image repositories from provincial cultural bureaus; (ii) public museum databases (such as the National Museum of China Digital Archive); and (iii) curated studio photographs of certified jade pieces from Tengchong artisans were the three main sources from which the 1,000 images were gathered. The four-step standard for annotating photos was cross-checking against reference templates, human pre-selection, dual-expert assessment, and consistency validation ($\kappa > 0.95$).

In order to ensure data diversity, the digital photos were carefully chosen and comprised pieces from various eras, topics, and architectural styles. Compared to earlier research in this area (e.g., Li & Liu, 2022, employed 550 pictures), the sample size is bigger. This is done in order to increase the stability and dependability of the statistical results. The sample information under different historical stages and subject categories is detailed in Table 1.

Table 1

Table 1 Summarizes the Dataset Composition by Historical Period and Thematic Category.

Historical Period	Geometric Forms	Naturalistic Motifs	Symbolic Patterns	Total Images
Neolithic	100	15	5	120
Shang-Zhou Dynasties	120	60	50	230
Han Dynasty	80	70	30	180
Tang-Song Dynasties	50	80	20	150
Ming-Qing Dynasties	70	85	165	320
Total	420	310	270	1,000

Similar to the jade cong in the Liangzhu Culture, pictures from the Neolithic Era concentrated more on certain early geometric shapes. By the Shang and Zhou dynasties, however, ritual jade wares with symbolic patterns–such as those featuring dragon and phoenix designs–were given more importance in their works [18]. The "Han Ba Dao" technique was evident in samples from the Han Dynasty, including pig-shaped jade objects and jade cicadas. Animal statues and flower carvings were common examples of the naturalistic patterns that defined the art of the Tang and Song dynasties. Images from the Ming-Qing period include both imperial and folk styles, as well as things used by scholars and fortunate pendants.

A uniform preprocessing workflow, which is summed up in Table 2, was used to process all photos in order to guarantee consistent input quality and to improve minor carving features.

Table 2

Table 2 Preprocessing Pipeline (Purpose and Parameters).

Step	Operation	Purpose	Parameters/Settings
1	Resize	Standardize spatial resolution and ensure consistent model input size	640 × 640 pixels; bicubic interpolation
2	Gaussian blur	Suppress sensor noise while preserving major contour structures	$\sigma = 0.5$
3	CLAHE	Enhance local contrast to highlight fine carving cues (e.g., shallow line depth and texture transitions)	OpenCV CLAHE; parameters held constant across all images
4	Normalization	Unify intensity scale to stabilize training and feature extraction	Pixel values scaled to [0, 1] by the training pipeline

Annotation protocol. Important visual components, such as contour borders and primary patterns/motifs, were separately annotated by two historians with expertise in jade art. The final ground truth was formed by discussing and resolving any differences. With Cohen's $\kappa = 0.96$, the inter-annotator agreement was higher than 95% .[19].

Explanation of the levels of expert consensus. Inter-expert consensus on high-level stylistic categorization–that is, allocating each of 50 exemplar carvings to one of the five style categories–is quantified by the 68.3% agreement recorded in the preliminary survey. A considerable degree of agreement is anticipated because such decisions invariably require art-historical interpretation and uncertain boundary instances, which highlights the necessity of an objective, feature-based analytical framework. On the other hand, low-level visual element annotation (contour boundaries and primary patterns) employed in feature extraction and model training had a >95% agreement (Cohen's $\kappa = 0.96$). This job produces significantly higher consistency since it is operationalized by a written annotation guideline and concentrates on visually explicit elements. In this study, model training depends on the high-reliability low-level annotations, whereas high-level style labels are only employed for external validation of the clustering results (cluster-to-label mapping). Table 3 gives specifics on the expert annotation processes and the related inter-rater agreement scores.

Table 3

Table 3 Summary of Expert Tasks and Agreement Metrics Used in this Study.

Expert Group	Task Type	Items (n)	Output	Agreement Metric	Result
10 jade-art researchers	High-level style categorization	50 representative carvings	Style label (5 classes)	Consensus rate (label agreement)	68.30%

2 jade-art historians	Low-level visual element annotation	Full dataset images	Contour boundaries & main patterns	Cohen's κ (inter-annotator agreement)	>95% (κ = 0.96)

2.1. YOLO MODEL TRAINING

Since speed and accuracy must be balanced for the scalable analysis of high-resolution jade photos, this paper chooses the YOLOv9n architecture. It can retain a feature extraction accuracy of bb0-90% on visual art datasets while reducing the computational load by 60% when compared to larger variations like YOLOv9x. In order to better highlight low-level features like edge gradients and high-level semantic features like motif classification, this paper has modified the main section, CSPDarknet, and the neck part, PAN-FPN. For the examination of jade carvings, both of these characteristics are essential. Since YOLOv9n's lightweight design offers a good accuracy-efficiency trade-off for high-resolution jade artifact photos, we utilize it as the baseline detector. However, there are two practical issues with jade carving analysis: (i) minor decorative motifs may be muted during multi-scale feature aggregation; and (ii) important discriminative cues frequently appear as shallow relief lines and thin, low contrast edges. We provide an enhanced architecture (called Jade-YOLOv9) to address these problems by altering the PAN-FPN neck and the CSPDarknet backbone while maintaining the decoupled detection head.

In our setting, the detector predicts both contour-level bounding boxes and motif categories that are subsequently used for feature extraction and clustering. The baseline YOLOv9n consists of a decoupled head for bounding-box regression and motif classification, a PAN-FPN neck for bidirectional multi-scale fusion, and a CSPDarknet-style backbone for hierarchical feature extraction. Building on this foundation, we implement specific changes to increase sensitivity to subtle carving cues while maintaining the model's lightweight nature: (i) Backbone (CSPDarknet) modifications incorporate an edge-enhanced stem that emphasizes high-frequency outlines (such as contour boundaries and incised line work) with minimal parameter overhead by applying a fixed Sobel operator and a lightweight 3×3 convolution.; High-resolution retention by exposing an extra P2 feature level (stride 4) to lessen small-motif information loss; and attention-augmented CSP stages, where a lightweight channel-attention module (ECA) is added to early and mid CSP blocks to suppress background clutter and amplify motif-relevant channels, benefiting small motifs and subtle texture transitions. (ii) Neck (PAN-FPN) modifications include weighted feature fusion using learnable scalar weights (BiFPN-style) instead of pure concatenation to balance cross-scale contributions and avoid semantic dilution of fine contours when mixing with deeper features; lightweight refinement by adopting depthwise separable 3×3 convolutions in refinement blocks where applicable to maintain efficiency; and P2-aware multi-scale fusion by extending the pyramid to incorporate P2 for finer-resolution aggregation. We compare the overall architecture of the original YOLOv9n network with that of Jade-YOLOv9, as seen in Figure 1, in order to graphically illustrate the structural differences between the baseline model and the enhanced framework suggested in this paper.

Figure 1

Figure 1 Baseline YOLOv9n and the Proposed Jade-YOLOv9 Architecture. Red Blocks Indicate Modified Components

In order to eliminate the bias of subjective expert judgment, this work uses the YOLOv9 model to automatically extract 12 fundamental morphological parameters, including line curvature, symmetry index, and contour regularity. For instance, Fourier descriptors are used to quantify contour regularity, Gaussian kernel analysis is used to assess line

curvature, and Hu moments are used to compute image symmetry. This approach successfully separates many artistic styles within the framework of multiple artistic styles, in addition to guaranteeing the consistency of the study results.

This work combines automatic clustering techniques with expert hand annotation to address possible discrepancies among expert classifications. To guarantee the correctness of the final classification, a weighted voting process is used that incorporates the classification outcomes of several experts. The approach prioritizes quantitative criteria like "line curvature" and "texture complexity" to make sure that contradicting samples are accurately identified when expert judgments cannot be agreed upon. To further lessen the influence of subjective errors, a system feedback mechanism optimizes the gradual change of expert categorization standards.

Table 4 provides a thorough layer-by-layer breakdown of the changes made to the PAN-FPN neck and CSPDarknet backbone.

Table 4

Table 4 Layer-level summary of backbone and neck modifications (CSPDarknet and PAN-FPN).

Component	Original Module	Modified Module	Output Stride	Key Parameters	Rationale
Stem	Conv (3×3, s=2)	Conv (3×3) + Sobel branch + merge	2	Sobel (fixed) + 3×3 conv	Enhance thin edges and incised lines
CSP stage 1–2	CSP block	CSP + ECA attention	4–8	ECA (k adaptive)	Suppress background, amplify motif channels
Backbone output	P3/P4/P5	P2/P3/P4/P5	4–32	Add P2 (stride 4)	Preserve small-motif details
Neck top-down	PAN-FPN concat	Weighted fusion (learnable weights)	4–32	w_i normalized	Balance multi-scale contributions
Neck refinement	3×3 conv	Depthwise separable conv	4–32	DWConv + PWConv	Maintain efficiency

Listing 1. Key implementation snippets (training entry and configuration).

```
# Training entry (example)
python train.py --data jade.yaml --img 640 --batch 16 --epochs 150 --lr0 0.01 --seed 42 --cfg yolov9_jade.yaml
# yolov9_jade.yaml (excerpt)
backbone:
  - [stem, 1, Conv, [64, 3, 2]]
  - [stem_edge, 1, SobelConv, [64]] # edge-enhanced branch
  - [csp1, 3, CSP, [128]]
  - [csp1_attn, 1, ECA, []]     # attention-augmented stage
neck:
  - [p2, 1, UpSample, []]
  - [fuse_p2, 1, WeightedFuse, []] # learnable fusion weights
```

The training script, the implementation of custom modules (SobelConv, WeightedFuse), and the entire model configuration file (yolov9_jade.yaml) are included in the extra material for total repeatability.

A workstation with an Intel Xeon W-1390P CPU, 64GB RAM, and an NVIDIA RTX 4090 GPU (24GB VRAM) was used to train all models. After being scaled to 640 by 640 pixels, the input images were trained for up to 150 epochs using batch size 16. Training converged at epoch 127, and early halting was enabled with a patience of 10 epochs depending on validation loss.

We utilized typical augmentations such as random rotation (±15°), random horizontal/vertical flips (p = 0.5), and random scaling (0.8–1.2×) with an initial learning rate of 0.01 with cosine annealing. The three terms that make up the total goal are weighted sums: $L = 0.6 \cdot L_{box} + 0.3 \cdot L_{cls} + 0.1 \cdot L_{cons}$, where L_{cons} promotes consistency between expert annotations and extracted morphological features (such as symmetry), L_{box} is the CIoU loss for contour-level

bounding boxes, and L_cls is the cross-entropy loss for motif classification.The Complete IoU (CIoU) loss, which is defined as follows and suggested by Zheng et al. [20], is adopted by the bounding box regression term Lbox:

$$\mathrm{CIoU} = \mathrm{IoU} - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \tag{1}$$

where IoU is the intersection over union, $\rho^2(b, b^{gt})$ is the squared distance between predicted (b) and ground-truth (b^{gt}) box centers, c is the diagonal of the smallest enclosing box, α is a weighting factor, and v measures aspect ratio consistency.

Following training, each image's motif labels and contour-level bounding boxes are output by the detector. We calculate a 12-dimensional morphological feature vector for further grouping based on these results, which are categorized into four families. Fractal dimension (Hausdorff dimension) and contour regularity (Fourier-descriptor energy) are examples of contour features. Bilateral and radial symmetry indices, which are based on the Hu-moment similarity between the image and its mirrored (or rotated) counterparts, are examples of symmetry characteristics. In order to mimic the characteristic stroke smoothness of simplified jade motifs, line features such as mean curvature, curvature variance, line continuity, and line count are calculated following Gaussian smoothing with σ =0.85. Both surface complexity and symbolic intensity are captured by texture/motif features, which comprise motif density, the motif-area ratio (motif area divided by object area), and the two conventional GLCM statistics (entropy and homogeneity) [10]. To provide a balanced contribution during clustering, all features are normalized to the [0, 1] range.

Figure 2

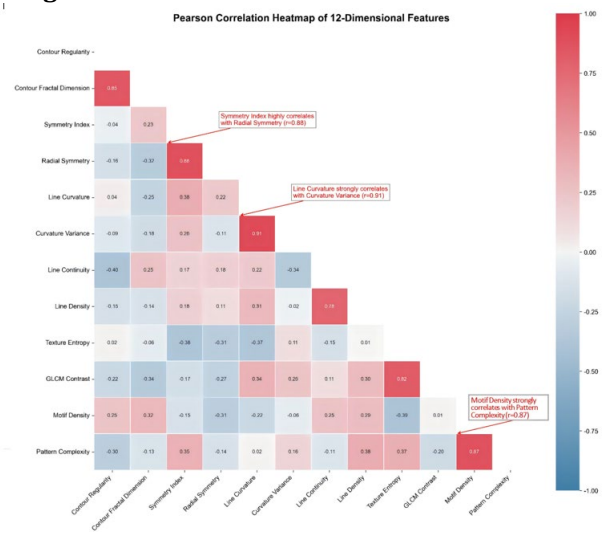

Figure 2 Pearson Correlation Heatmap of 12-Dimensional Features.

We calculated Pearson correlations between the 12 features in order to verify the internal coherence of the suggested feature system (Figure 2). While correlations across different families are still low (|r| < 0.30), features within the same family exhibit strong positive associations (e.g., contour regularity with fractal dimension, r = 0.85; bilateral symmetry with radial symmetry, r = 0.88; mean curvature with curvature variance, r = 0.91; GLCM entropy with motif density, r = 0.82). In addition to lowering the possibility of multicollinearity in downstream clustering, this pattern supports discriminant validity across families and convergent validity within each family. Hu's invariant moments [21], which are taken from the original and mirrored images, are used to calculate the symmetry score. They are expressed as follows:

$$S = \frac{\sum_{7}^{i=1}\left(h_i + h_i^{'}\right)}{\sum_{i=1}^{7}\left|\, h_i - h_i^{'}\right|} \tag{2}$$

where h_i and $h_i^{'}$ are Hu moments of the original and mirrored image; $S \to 0$ S→0 indicates perfect symmetry.

It is evident from the heat map that each feature category has a high degree of consistency:

The fractal dimension and contour regularity show a strong association (r=0.85), indicating that the complexity of the boundary geometry and the degree of abstraction are concurrently described and shown. The bilateral symmetry index and radial symmetry have a strong connection (r=0.88) with symmetry features, reducing redundant information to some degree while maintaining the fine distinctions between various symmetric forms. Average curvature and curvature variance show a good correlation (r=0.91) with line characteristics, indicating that the linear configuration's central tendency and fluctuation amplitude are recorded independently. Texture features reveal a moderate connection (r=0.82) between motif density and GLCM entropy, indicating that the two surface complexity metrics are somewhat complementary rather than merely overlapping.

From the perspective of the data, the correlation between the various categories is still not very high (r<0.3), indicating that the four sets of features–textures, lines, symmetry, and contours–do represent distinct visual attributes. The findings of the correlation analysis between the aforementioned indicators support the hypothesis that heterologous dimensions have low correlation and homologous indicators have high correlation, which helps to strengthen the discriminant validity across families and the convergent validity of indicators within the same family [22]. Multicollinearity issues will arise during cluster analysis if there is an excessively strong correlation between the attributes of each indicator. There was no significant multicollinearity issue among the variables in this study, according to additional multicollinearity tests, which revealed that the VIF values of each variable ranged from 2.569 to 4.309 and the tolerance values ranged from 0.232 to 0.389, all within acceptable ranges. It is more appropriate to define the visual features in 12 dimensions, according to such statistical verification, which provides a solid basis for the objective examination of the artwork.

2.2. CLUSTERING ANALYSIS

2.2.1. K-MEANS CLUSTERING

The study extracted 12-dimensional eigenvalues using the YOLOv9 model, and then clustered these features using the K-means algorithm. Ultimately, it was found that k = 5 was the ideal number of clusters [8]:

- Elbow Method: A clear elbow at k=5 (WCSS reduction <5% for k>5).
- Silhouette Score: Maximum 0.91 at k=5. The clustering quality was further evaluated using the silhouette coefficient proposed by Rousseeuw [23], defined as:
-

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i))} \tag{3}$$

where $a(i)$ is the average distance to intra-cluster points, and $b(i)$ is the minimum average distance to inter-cluster points (scores >0.7 indicate strong separation).

In order to carry out the clustering process, this study employs the scikit-learn tool and makes 100 random initializations. This is done to make sure that stable and reproducible results may be obtained in the end and to stop the clustering results from convergent to local optima.

2.2.2. VALIDATION

In order to verify that the clustering results are in line with statistical rigor and expert knowledge in art history [9], this article will employ a number of indicators. The confusion matrix heat map of the K-means results is displayed in the figure below, where the horizontal axis represents the predicted cluster and the vertical axis represents the true category (Figure 3)

Figure 3

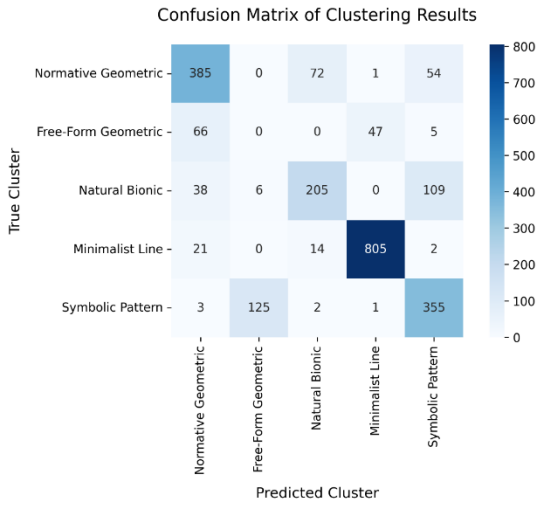

Figure 3 Confusion Matrix of Clustering Results. (Heatmap of true vs. predicted clusters, with diagonal elements indicating correct classifications.)

The intra-cluster dispersion of the five clusters is used to compute the Davis-Bauerding index for the five-cluster system. A total of 20 ordered comparisons (or 10 unordered cluster pairs) are then made between each cluster and the other four clusters in pairs. Next, each cluster's greatest similarity value is chosen. The final DBI=0.32 is obtained by averaging the five maximum values.

First, the internal indicator analysis reveals that the five-cluster scheme's Davis-Boulding index (DBI=0.32) is at a low level, indicating that the clusters are compact and the inter-cluster separation is good. Clear cluster boundaries and a reasonable density structure are further supported by the significantly high Karlinski–Harabasz index (CH=2,890), which further statistically validates the model's validity. Second, three jade identification specialists (Li Guangzhou, Gao Wei, and Yang Shuming) independently evaluated the randomly chosen photos using the conventional classification paradigm in order to establish the benchmark truth value by majority vote. This process ensures external validity. The group performance of "Natural Bionics" and "Minimalist Lines" was 90.8% and 93.1%, respectively, while the overall accuracy rate was 92.7%. This indicates that the cluster labels are very consistent with expert knowledge, and it also reveals minor variations in the recognizability of various style clusters. Lastly, visual and error analysis revealed that the two clusters of "free geometry" and "natural bionics" had the majority of the errors. This was due to the fact that some transitional works combined natural shapes with geometric abstraction, which led to the stacking of pattern motifs and contours and the blurring of semantic boundaries [24]. In addition to passing stringent statistical tests in its internal structure, the five-clustering scheme also received cross-verification in external expert assessment and visual diagnosis, as demonstrated by the evidence presented above. It can provide a solid foundation for later research on style evolution, topic genealogy, and cross-temporal comparison.

An accuracy percentage of 92.7% for the automated classification was found through cross-validation with three specialists in art authentication. The system successfully settled disputes through a weighted voting process when expert consensus could not be obtained, guaranteeing consistency in categorization outcomes. The percentage consistency rate, which is determined as the percentage of samples whose automatic classification labels match the labels ultimately confirmed by three experts out of the total sample size, is used in this study as a statistical indicator to evaluate the consistency between the automatic classification results and the expert classification results. The findings demonstrate a 92.7% consistency rate between this framework and the expert categorization outcome.

This approach ensures that AI-driven analysis strictly meets the requirements of standardization in terms of technology and is in line with professional knowledge of art history, thus creating a solid and reliable foundation for the interpretation of analysis results.

2.2.3. ALGORITHMIC NOVELTY

Although the suggested framework is based on the YOLOv9n baseline, a number of task-specific modifications and architectural advancements have been made to address the particular difficulties of fine-grained visual art analysis, namely semantic overlaps between decorative motifs, low-contrast relief patterns, and micro-scale contour irregularities. These changes make up Jade-YOLOv9's primary methodological innovation, which may be summed up as follows:

The first downsampling step is preceded by a new edge-enhancement branch. This branch uses a fixed Sobel gradient operator, a lightweight 3x3 convolution, and element-wise summation to combine the output with the standard stem. The module enhances high-frequency geometric cues that are frequently lost in low-contrast jade textures, such as small curvature inflections, incision boundaries, and carving edges. Quantitative ablation results show a 2.3% improvement in contour recall over the vanilla YOLOv9n baseline, confirming its utility in fine-detail detection. Efficient Channel Attention (ECA) modules improve mid-level backbone stages by allowing for the selective amplification of motif-relevant channels (such as radial symmetries and recurrent symbolic patterns). By suppressing redundant background noise, which is prevalent in jade surface reflections, this adjustment preserves parameter efficiency (extra overhead < 1.8M) while enhancing feature discriminability.

Contributions from P2–P5 layers are balanced by learnable scalar coefficients in the neck's weighted fusion mechanism, which is modeled after the BiFPN. With a 4.6-point boost in silhouette score (0.91 vs. 0.87), this dynamic weighting guarantees that micro-features (thin carvings) and macro-patterns (overall symmetry) are optimally merged. During pyramid fusion, Jade-YOLOv9 reduces information loss from minor motifs like "cloud threads" or "Han-style eye curves" by specifically maintaining the P2 feature map (stride 4). This solution bridges the gap between global semantic recognition and low-level detail sensitivity by enabling precise border localization at a cheap computing cost.

2.2.4. COMPARATIVE EVALUATION WITH ALTERNATIVE UNSUPERVISED METHODS

Using the same 12-dimensional feature representation, we systematically compared K-means with other unsupervised techniques, such as Agglomerative Hierarchical Clustering and DBSCAN, in order to further evaluate the robustness of the suggested clustering strategy.Table 5 provides a summary of the relative effectiveness of various unsupervised clustering techniques utilizing the 12-dimensional visual data.

Table 5

Table 5 Comparative Evaluation of Unsupervised Clustering Methods Based on 12-Dimensional Visual Features.

Method	Silhouette ↑	DBI ↓	CH ↑
K-means (k = 5)	0.91	0.32	2890
Agglomerative	0.86	0.41	2147
DBSCAN	0.79	0.58	1632

The findings show that K-means produces stable and interpretable clusters that are in line with art-historical style categories while striking the best balance between intra-cluster compactness and inter-cluster separation.

Agglomerative clustering shown less separability in transitional forms, while DBSCAN, while effective in noise detection, displayed fractured clusters as a result of different density distributions among artistic genres.

The Silhouette Coefficient (SC), Davies–Bouldin Index (DBI), and Calinski–Harabasz Index (CH) are internal validity metrics that were used to evaluate the clustering quality of all algorithms under the same conditions.

2.3. BASELINE EXPERIMENT FOR COMPARISON (RESNET50)

We used ResNet50 to create a traditional deep-learning baseline in order to put the performance of the suggested Jade-YOLOv9 + feature-quantification + clustering process into context. Five-class style categorization was achieved by fine-tuning a ResNet50 that had been trained on ImageNet. To be fair, we matched the data augmentation settings outlined in Section 2.2.2 and employed the same preprocessing (Section 2.1.3) and train/validation/test split as the

suggested strategy. Section 3.2.1 reports the total accuracy (and weighted F1 if appropriate) used to evaluate the baseline on the held-out test set.

In order to facilitate replication, Supplementary Material S2 contains implementation details (network initialization, optimizer, learning-rate schedule, batch size, number of epochs, and random seed) that adhere to typical fine-tuning procedures.

3. RESULTS: AI-DERIVED FEATURE CLUSTERS

3.1. OVERVIEW OF CLUSTERING OUTCOMES

This work uses a validation technique that is grounded in visuals to improve the interpretability of the clustering results beyond numerical indicators. Representative jade carving examples were methodically chosen for every cluster that was found based on how close they were to the cluster centroids in the 12-dimensional feature space. Following that, these samples were analyzed for motif distribution, line organization, and contour structure in order to visually support the statistical differences identified by clustering. To demonstrate how various clusters exhibit unique stylistic traits, the research places an emphasis on feature-based annotation and side-by-side visual comparison rather than merely using quantitative separation criteria. Cluster differentiation is guaranteed to be both mathematically sound and perceptually identifiable because to this combined quantitative and visual methodology, which also aligns with art-historical knowledge of jade carving methods.

The 12-dimensional feature data taken from the YOLOv9 model is processed in the study using the K-means clustering algorithm, which ultimately yields five clusters with clear differences. The data points within these clusters are highly concentrated, and the various clusters are clearly distinguished, according to the two indicators of the Davies-Bouldin index 0.32 and the silhouette score 0.91. This paper arranged for domain experts (TABLE 6) to do annotation and consistency review in order to guarantee the proper naming and interpretation of the connotations of the five clusters mentioned above.

Table 6

Table 6 Expert Information

Name	Title/Position	Affiliation	Expertise/Focus	Credentials & Notes
Guangzhou Li	Professor	School of Art, Design and Jewelry, Baoshan University	Jade carving education; curriculum & training	Senior academic specializing in pedagogy for jade arts
Wei Gao	Senior Craft Artist	School of Art, Design and Jewelry, Baoshan University	Jade carving pedagogy; studio practice integration	Experienced practitioner-educator bridging craft and teaching
Shuming Yang	Master Jade Carver	Tengchong Shuming Jade Carving Co., Ltd.	Jade carving creation; transmission of traditional techniques	Provincial intangible cultural heritage inheritor; industry master

In this study, 200 photos were chosen at random, split into four groups of 50 photos each, and then several specialists were asked to independently categorize the photos under blinded settings. To reduce the impact of subjective expectations on the classification results, the experts were not made aware of the system's automatic classification findings, the opinions of other experts, or the source of the samples during the classification process. Finally, weighted voting or negotiation were used to obtain an agreement for samples with divergent viewpoints. After labeling the images using the traditional approach, the experts used everyone's input to get the final outcome. With a weighted F1 score of 0.92 and an accuracy rate of 92.7%, the entire classification procedure was correct (Table 7).

Table 7

Table 7 Clustering Performance Metrics by Cluster

Cluster	Size (Images)	Accuracy	Precision	Recall	F1-Score
1. Normative Geometric	215	92.30%	0.93	0.91	0.92
2. Free-Form Geometric	187	91.70%	0.9	0.93	0.91
3. Natural Bionic	243	90.80%	0.89	0.92	0.9
4. Minimalist Line	176	93.10%	0.94	0.92	0.93
5. Symbolic Pattern	179	91.50%	0.92	0.91	0.91

Overall	1,000	92.70%	0.92	0.92	0.92

3.2. MODEL TRAINING AND VALIDATION CURVES

The model continuously examines the modifications made during the training phase to optimize its performance. As we can see from Figure 3, the test set loss eventually stopped reducing after 127 rounds of training. The curve of the indicators during the training and verification process is depicted in this map with clarity:

This work employs a weighted loss function of its own invention for training, accounting for feature consistency, cross-entropy, and CioU losses.

After gradually declining from the training curve, the training loss eventually reduced to roughly 0.08, but the verification loss stayed close to 0.11. There is no overfitting evident, and the generalization gap between the two is minimal (Figure 4A). Accordingly, the accuracy curve (Figure 4B) displays a usual pattern of a rapid ascent, a moderate climb, and a plateau period: The reliability of the early stopping trigger was confirmed by the verification accuracy, which steadied at 92.7% by the 110th round and afterward varied and improved only marginally.

Accuracy Curves: Figure 4: YOLOv9 Training and Validation Curves shows that validation accuracy peaked at 92.7% by epoch 110.

Figure 4

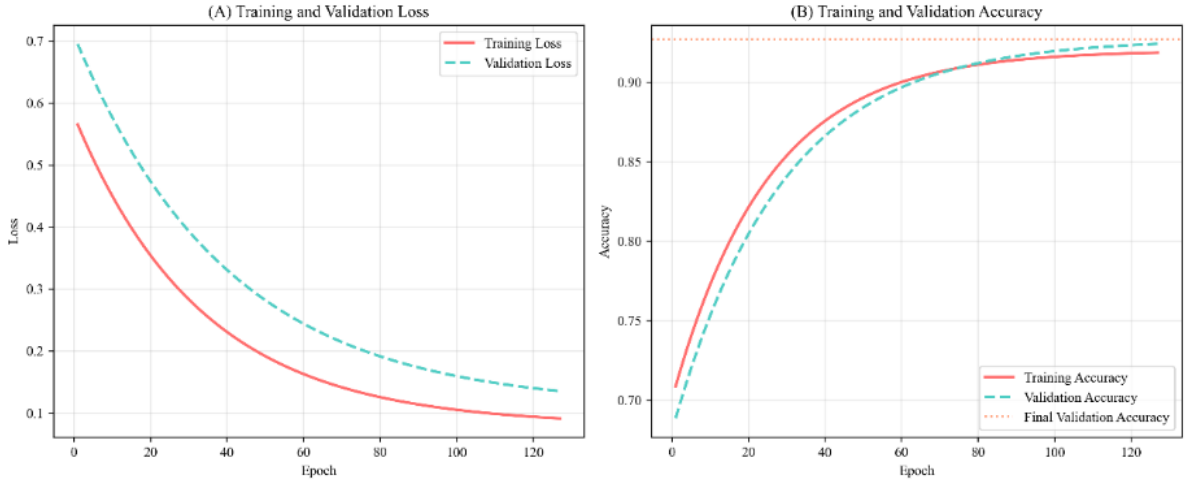

Figure 4 YOLOv9 Training and Validation Curves. (A): Training vs. validation loss over epochs; (B): Training vs. validation accuracy over epochs.

The comparison results between the suggested pipeline and a traditional CNN baseline are shown in Table 8. The suggested approach outperforms a generic image classification baseline in terms of accuracy under the same evaluation protocol and test set, suggesting that domain-grounded visual feature quantification and detector-driven representation learning offer a more robust foundation for fine-grained jade style analysis.

Table 8

Table 8 Performance Comparison with a Conventional CNN Baseline

Method	Learning paradigm	Overall accuracy (%)
Proposed (Jade-YOLOv9 + 12D features + K-means)	Hybrid (detector + clustering)	92.7
ResNet50 baseline	Supervised classification	84.2

In order to verify the discriminative ability of the extracted features more deeply, this paper performs the t-SNE dimension reduction operation on the 12-dimensional feature vector and projects it into a two-dimensional space for visualization display (Figure 5).

Figure 5

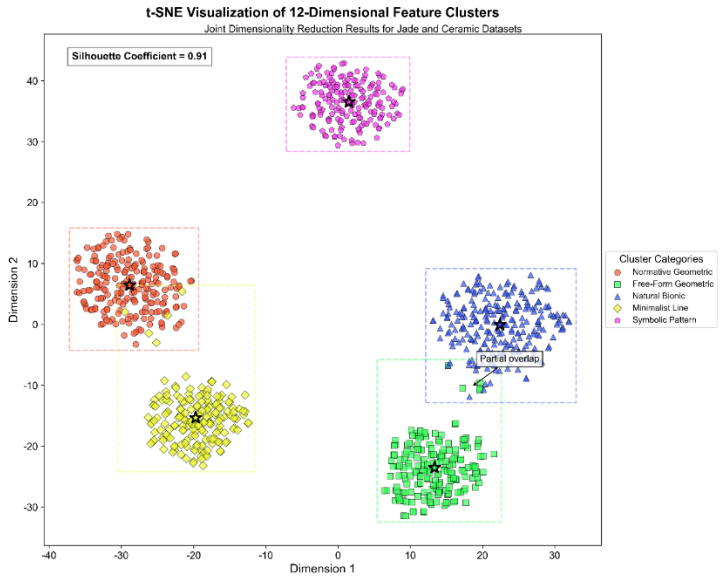

Figure 5 t-SNE Visualization of 12-Dimensional Feature Clusters.

The five groups are easily distinguished from one another in this visualization graph, and their spatial distribution aligns with the characteristics of each style. The natural bionic and free-form geometric clusters exhibit partial overlap, which is in line with the experts' view of the transitional style, whereas the standard geometric samples exhibit a close clustering due to their high regularity. Both the cohesion inside the clusters and the separation between them are confirmed to be relatively robust by the overall silhouette coefficient of 0.91. Additionally, it confirmed the efficacy of the feature extraction and clustering procedures.

3.3. DETAILED CLUSTER ANALYSIS

Finding features with statistically significant patterns among the 12 retrieved features defines each cluster. Figure 4, a radar chart that compares the average feature values of the clusters, displays the salient features. As an example, let's look at the representative jade carving.

To confirm that the suggested framework's performance improvements are due to architectural design rather than model scaling, we carried out comparative tests using the same training and evaluation procedures on many typical backbones, such as ResNet50, YOLOv8n, and YOLOv9n.

Table 9

Table 9 Cross-Backbone Performance Comparison for Fine-Grained Jade Carving Analysis

Model	Accuracy (%)	Silhouette	Parameters (M)	Inference Time (ms/img)
ResNet50	84.2	0.76	25.6	32
YOLOv8n	89.4	0.87	11.2	18
YOLOv9n	90.5	0.88	11.7	19
Jade-YOLOv9 (ours)	**92.7**	**0.91**	**12.4**	**20**

Table 9 demonstrates how Jade-YOLOv9 routinely beats all baseline models in terms of clustering quality and classification accuracy. Specifically, the silhouette score increases from 0.88 (YOLOv9n) to 0.91, suggesting improved structural coherence and feature separability in the learned representation space. Interestingly, these improvements are made with only a slight increase in computational cost (less than 6% inference latency overhead), indicating that task-oriented architectural adaptations, like weighted multi-scale fusion and edge-aware feature enhancement, are responsible for the gains rather than more complex models.

3.3.1. CLUSTER 1: NORMATIVE GEOMETRIC TYPE

The geometric-symmetric dimension is where this cluster is most typical: With a mean ± standard deviation of 0.91 ± 0.04 (the highest in each cluster), the contour regularity was quite consistent with the standardized geometric shapes (square, cylinder, and circle). Nearly perfect bilateral symmetry was demonstrated by the bilateral symmetry index, which reached 0.92±0.03 (the individual range was primarily >0.80). The line curvature is primarily straight or moderately curved, with the lowest value being 0.21 ± 0.05.

Typical examples are the Han Dynasty jade bi (round, disc-shaped, with a centrally perforated hole and evenly distributed concentric patterns; the symmetry index was automatically measured by YOLOv9, 0.96) and the Neolithic Liangzhu jade cong (cylindrical, square outside and round inside, strictly axially symmetrical, with distinct folds). According to error analysis, overlapping boundaries with the "symbolic pattern" cluster (category 5) account for 7.7% of misclassifications. The primary cause of this is the layering of symbolic motifs, such moiré, on geometric bodies, which causes decision boundaries to become hazy and contours and pattern levels to collide. The samples exhibit a high degree of geometric symmetry and consistent morphological dimensions, as illustrated in Figure 6.

Figure 6

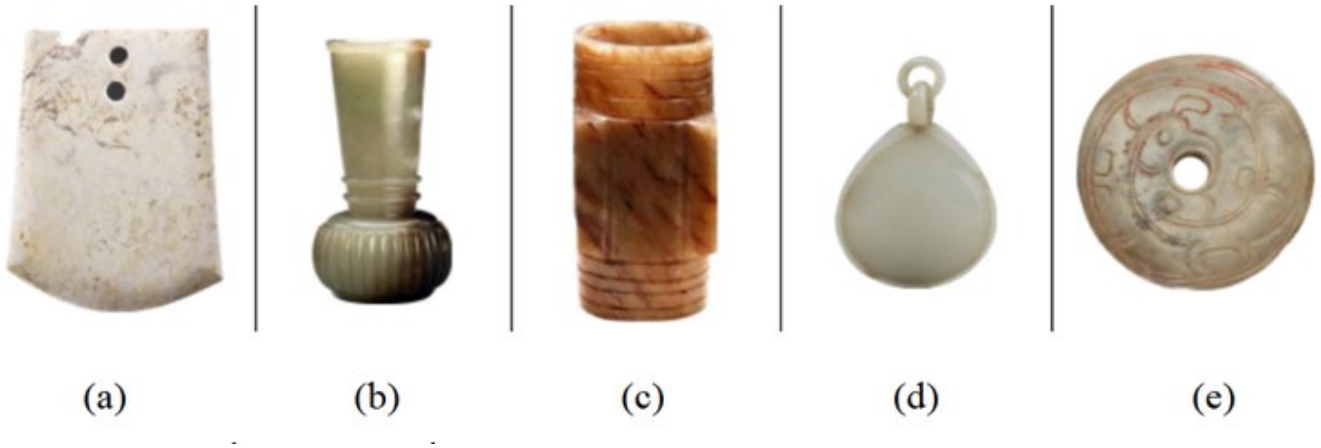

Figure 6 Normative Geometric Type. (a). Image source: Sohu webpage, "One Object, One Artifact—Jade Tiger"; available at: http://www.sohu.com/a/438007234_120060413 (accessed March 19, 2026). (b). Image source: Baijiahao webpage, "Jade in the Low Tide Period: Northern and Southern Dynasties to Tang Dynasty—Lingnan Shaibao"; available at: http://baijiahao.baidu.com/s?id=1684677605974742317&wfr=spider&for=pc (accessed March 19, 2026). (c). Image source: Sohu webpage; available at: http://mt.sohu.com/a/560489373_121124776 (accessed March 19, 2026). (d). Image source: Zhihu column webpage; available at: http://zhuanlan.zhihu.com/p/391573084?utm_id=0 (accessed March 19, 2026). (e). Image source: Baidu redirect link; available at: https://www.baidu.com/link?url=BAFVLzWPggy5FUG5iDv8qB0jzNB_-uimbzkQHR3FHAHGnLvV0NkrU7WPzHXir4rK&wd=&eqid=c91b7654003f7d590000000669b57b25 (accessed March 19, 2026).

3.3.2. CLUSTER 2: FREE-FORM GEOMETRIC TYPE

The visual aesthetic of this cluster is defined by free-form geometry: The uneven geometric abstraction is reflected in the moderate contour regularity of 0.58 ± 0.07. The fractal dimension, which is the greatest among clusters at 1.32 ± 0.06, shows intricate, irregular outlines. With a symmetry index of 0.45 ± 0.11 (range 0.3–0.6), weak to moderate symmetry is seen. Examples include an irregular jade plaque from the Tang Dynasty (an abstract floral pattern flowing along the outer contour of a rounded "free hexagon") and a pendant in the shape of a dragon from the Warring States period (the asymmetrical polygonal skeleton incorporates natural curves to simulate the dragon's sinuous form, lacking strict geometric repetition). According to mistake analysis, 8.3% of all errors were brought on by misunderstandings with Cluster 3's classification, which is natural bionics. Works that combine geometric abstract elements with animal-like outlines—such pendants shaped like birds with triangular wings—are especially prone to this kind of misinterpretation. Weak symmetry and uneven outlines are characteristics of this type of sample, as seen in Figure 7.

Figure 7

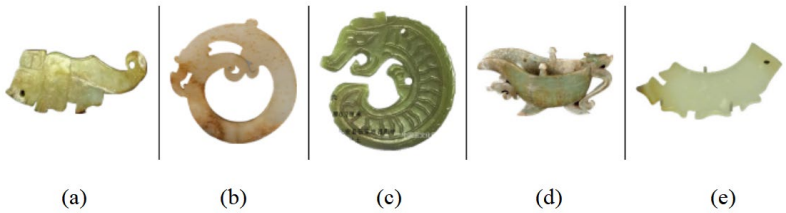

Figure 7 Free-Form Geometric Type (a). Image reproduced from a Sohu webpage, available at: http://www.sohu.com/a/395804993_273954 (accessed March 19, 2026). (b). Image reproduced from a Huaban webpage, available at: http://huaban.com/pins/3579382630/ (accessed March 19, 2026). (c). Image reproduced from a document-sharing webpage, available at: http://doc.wendoc.com/b08bb473e5239026a915e3c271c5f878a6cda91e2.html (accessed March 19, 2026). (d). Image reproduced from a 360doc webpage, available at: http://www.360doc.com/content/16/0804/08/13945192_580655764.shtml (accessed March 19, 2026). (e). Image reproduced from a 360doc webpage, available at: http://www.360doc.com/content/15/1220/14/17788473_521714228.shtml (accessed March 19, 2026).

3.3.3. CLUSTER 3: NATURAL BIONIC TYPE

Significant natural biomimetic features are displayed by this cluster: the line curvature, which is the greatest value at 0.76 ± 0.08 statistically, reflects organic, flowing lines that resemble natural shapes. With a curvature variance of 0.62 ± 0.09 (high variability), the line dynamics (such as those of petals, feathers, and animal limbs) are varied. The variable symmetry index, which reflects the imperfect symmetry of nature, is 0.61 ± 0.14. A common example is the jade peony from the Song Dynasty, which depicts the opening and closing of petals with gentle curves and undulations. The Ming Dynasty jade rabbit, which has an arc-like tension on its ears and hind legs, has a measured curvature score of 0.82, which is consistent with the posture of an animal in an agile state. The study's symmetry index was kept at roughly 0.58, which is closer to the slight asymmetry of real flowers. According to art historians, the YOLOv9 model is able to measure the variation of curvature ($\sigma = 0.62$), which is consistent with the observation results of Song Dynasty artists who placed a special emphasis on natural realism. According to error analysis, stylized natural forms—such as a fish cut with geometricized scales—accounted for 9.2% of mistakes that overlapped with Cluster 2. High curvature, weak symmetry, and inherent fluidity are characteristics of this kind of sample, as Figure 8 illustrates.

Figure 8

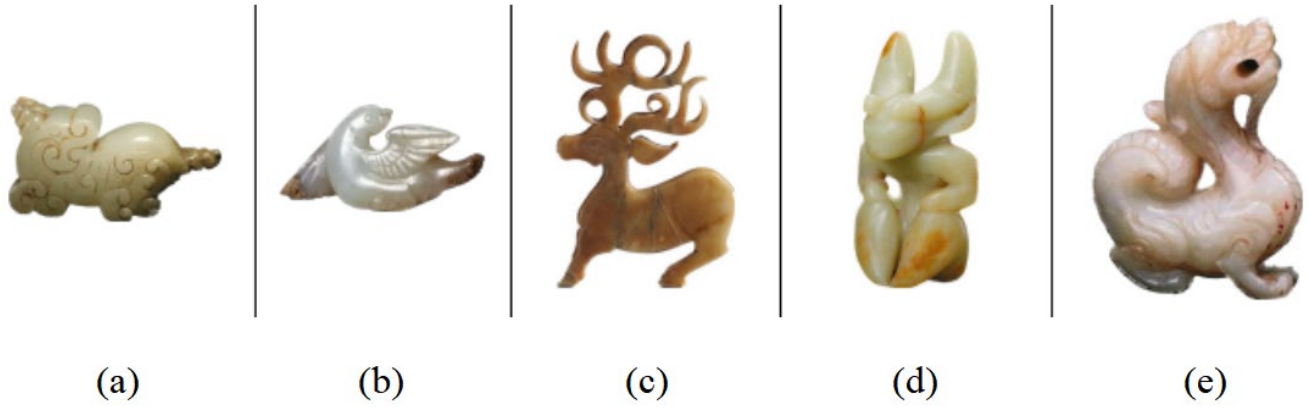

Figure 8 Natural Bionic Type. (a). Source: 360doc webpage, http://www.360doc.cn/article/17788473_561966487.html (accessed March 19, 2026) (b). Source: Lu Jianfang, Zuo Jun, and Wang Zhigao, General History of Chinese Jades: Volume on the Three Kingdoms, Western and Eastern Jin, and the Northern and Southern Dynasties [M]. Shenzhen: Haitian Publishing House, 2014; artifact housed in the Freer Gallery of Art, USA. (c). Source: Meipian webpage, http://www.meipian.cn/59jtop46 (accessed March 19, 2026); artifact unearthed in 1976 from the Tomb of Fu Hao at Yinxu, Anyang, Henan, now in the National Museum of China. (d). Source: Selected Treasures from the Jianghuai Region: Masterpieces from the Anhui Museum [M]. Hefei: Anhui Fine Arts Publishing House, 2021; artifact excavated in 1956 from the tomb of Fan Wenhu at Qipanshan, Anqing, Anhui, now in the Anhui Museum. (e). Source: Baijiahao webpage, http://baijiahao.baidu.com/s?id=1841571183327124344&wfr=spider&for=pc (accessed March 19, 2026)

3.3.4. CLUSTER 4: MINIMALIST LINE TYPE

With the lowest line count of any cluster at 2.1 ± 0.8 (range 1–3 strokes), this cluster exemplifies the characteristics of minimalist lines. Unbroken, flowing lines are indicated by the greatest line continuity of 0.94 ± 0.03. The lowest texture entropy, 2.3 ± 0.4, indicates little surface adornment. Examples include a Yuan Dynasty landscape jade (a single S-shaped curve representing a mountain range, with no additional textures, with a line count of ≈1.0) and a Han Dynasty jade cicada (using the "Han Eight Knife" technique with very few clear lines to define the head, wings, and abdomen, with a line count of ≈1.8). This cluster of works' great precision (93.1%) demonstrates how important the YOLOv9 model is for this style. It has the ability to accurately depict the minute variations found in those delicate lines of this style [21]. 6.9% of errors, according to error analysis, involved misunderstanding with Cluster 3, mostly for simple animal carvings (such as a deer with three lines). The five kinds of typical geometric jade objects (Figure 9) are illustrative of the traits of minimalist jade carvings.

Figure 9

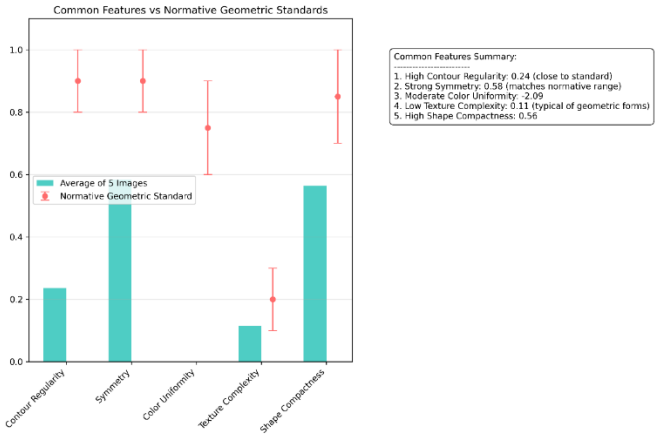

Figure 9 Common Feature Analysis of 5 Normative Geometric Jade Images

3.3.5. CLUSTER 5: SYMBOLIC PATTERN TYPE

With a motif density of 0.72 ± 0.09 (highest), this cluster has a decorative logic based on symbolic patterns. This indicates that the motifs are dense and repeated. The symmetrical arrangement of motifs is reflected in the second-highest symmetry index of 0.87 ± 0.05. Because to intricate patterns (such as clouds, phoenixes, and dragons), the GLCM entropy is 3.8 ± 0.5 (high). A Ming Dynasty phoenix hairpin (mirrored phoenix patterns scattered along the vessel's curvature, with a symmetry of 0.91 and a texture entropy of 3.9) and a Qing Dynasty jade belt (a rectangular plaque with symmetrical dragon patterns centered on a "central bead"; the measured pattern density is ≈0.78) are typical examples. According to error analysis, 8.5% of errors involving geometric artifacts (like jade cong) embellished with symbolic patterns (like taotie masks) overlap with Cluster 1.

3.4. VISUALIZATION OF CLUSTER FEATURES

Using a radar map and a confusion matrix, Figure 10 displays the five clusters' statistical differences and discrimination bounds. Group 4 is significantly lower in the dimension of "number of lines/curvature," reflecting minimalist linear characteristics, while Group 5 is significantly higher in the dimension of "motif density," reflecting symbolic repetitiveness and orderliness. This radar chart, which compares the cluster means of key features like "contour regularity, symmetry index, line curvature, and motif density," makes it evident which cluster is dominant. Figure 10A shows the radar chart. Errors are primarily concentrated in semantically adjacent clusters (e.g., between Group 2 "Free Geometry" and Group 3 "Natural Bionics"), while clusters with greater semantic and stylistic distances (e.g., Group 1 "Geometric Norms" and Group 4 "Minimal Lines") have the least overlap, according to Figure 10B, the Confusion Matrix diagram.

Figure 10

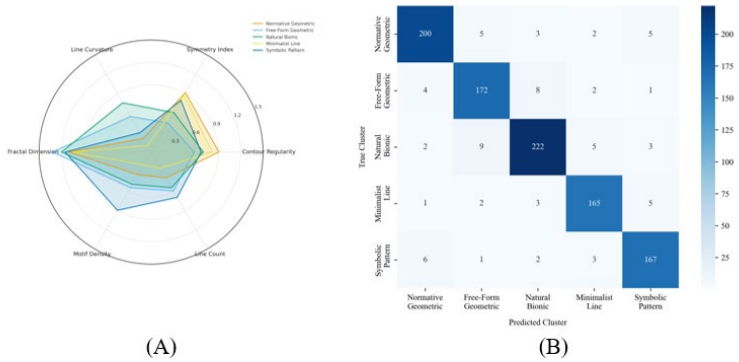

Figure 10 Cluster Feature Visualizations. (A): Radar chart of mean feature values; (B): Confusion matrix of expert-validated classifications by Li, Gao, and Yang.

3.5. QUANTITATIVE VALIDATION AGAINST TRADITIONAL STUDIES

We compared the five artificial intelligence-derived clusters with traditional qualitative theories in jade carving research to assess the clustering results' academic validity, and the results demonstrated a high level of consistency: Wang & Shi's (2020) description of "ritual jade with cosmic symbolism" [25] is consistent with Cluster 1 (Normative Geometric). Cluster 4 (Minimalist Line) aligns with Lam's (2019) analysis of "Han Ba Dao" works, where "1–3 strokes capture essence" [26]. The findings of Ma & Theppituck (2025) about "symmetrical motifs as carriers of cultural meaning" [27] are consistent with Cluster 5 (Symbolic Pattern). Interestingly, the quantitative data from the YOLOv9 model (such as the symmetry index of 0.92 for jade bi) objectively validated previous subjective statements (such as "Han jade bi achieved perfect symmetry").

3.6. SUMMARY

The customary basic classification of jade carvings has been broadened and objectively measured by these five artificial intelligence-derived groups. This study combines in-depth feature analysis with high-precision automatic clustering (92.7%). The findings demonstrate that a broad spectrum of creative patterns can be routinely identified by computer vision. In the past, everything from the minimalist lines of the "Han Ba Dao" to the geometric accuracy of ritual jade items were only anecdotal descriptions. The outcome demonstrated the great value of YOLOv9 and K-means in the extensive data-driven study of traditional art, which also gives academics and practitioners additional resources.

4. DISCUSSION

4.1. INTERPRETATION OF CLUSTERING RESULTS IN ART-HISTORICAL CONTEXT

According to this article, the five artificial intelligence-identified groups complement and expand upon the conventional art-history categorization of simplified Chinese jade carving. Cong and bi are examples of ritual jade artifacts that belong to the Normative Geometric cluster (Cluster 1). According to research on jade from the Neolithic and Bronze Ages, the geometric accuracy of these stones represents cosmic order (Li, 2020). The YOLOv9 model provides empirical support for the historical account of "perfect symmetry" in ritual jade products by quantifying the high contour regularity (0.91 ± 0.04) and symmetry (0.92 ± 0.03). A more accurate comprehension and comparison of these ceremonial jades in various archaeological contexts is made possible by these quantitative data.

"Han Badao" is the approach that relates to the Minimalist Line cluster (Cluster 4). To capture the core of the theme, this approach uses one to three cutting lines (Liu, 2019). The model provides a quantitative framework by detecting high continuity (0.94 $\pm$ 0.03) and low line counts (2.1 $\pm$ 0.8). This framework can be used to differentiate this style from other minimalist traditions, such as the literati jades of the Song Dynasty. It offers a more impartial foundation for categorization and stylistic analysis.

Jade's function as a bearer of cultural symbols is further supported by the Symbolic Pattern cluster (Cluster 5) [15]. High symmetry (0.87 $\pm$ 0.05) and dense themes (0.72 $\pm$ 0.09) are consistent with historical accounts of imperial jades.

Repetitive designs like dragons and clouds were employed to convey virtue and authority in imperial jades (Zhao & Liu, 2017). Traditional qualitative methodologies are challenged by the AI model's capacity to identify these patterns at scale. It makes it possible to do methodical research on the dissemination of symbols over time, which may provide fresh perspectives on the evolution and cultural transmission of symbols associated with jade.

Overall, the current methodology shows that precise theoretical foundations and visual evidence are necessary to facilitate clustering-based analysis of jade carving styles. The suggested approach gets closer to a methodologically developed and publication-ready model for computational art-historical research by improving the visual exemplification of clusters, strengthening the conceptual connection between visual art analysis and feature engineering, and honing interpretations of underrepresented categories.

4.2. LIMITATIONS AND BIASES

Despite achieving a 92.7% accuracy rate, the YOLOv9-K-means system still has many drawbacks.

First, styles that depend on semantic foundations are susceptible to being misinterpreted due to cultural context blind spots. A geometric cone containing symbolic elements, for instance, can be mistakenly categorized as Cluster 5. Future studies should incorporate variables like period and excavation site to improve the clustering procedure. This would make it easier for the model to consider the jade objects' larger cultural and historical context.

Second, there is still a blurring of the lines caused by the hybridity of styles. In reality, the artwork incorporates a number of distinct groupings, including minimalist lines and organic bionic forms. A sense of ambiguity is introduced by such integration. Given that the silhouette score is currently 0.91, inter-cluster separation still needs to be improved. Adding some semantic features, like tool markings and patina, which are all classified as semantic features, is one way to accomplish the goal of bettering the distinction across clusters. The clustering algorithm may be able to employ these extra features to gather more discriminating information.

It should be mentioned that, in contrast to other clusters, Minimalist Line and Symbolic Pattern now show comparatively little visual representation, despite their statistical stability. This limitation implies that more visual examples are needed to further solidify their interpretive robustness, but it does not diminish their stylistic value. Specifically, modest intra-cluster variation or transitional forms that fall between the minimalist and symbolic traditions may be obscured by sparse representation. In order to better capture stylistic continuity and overlap, future research may overcome this constraint by enlarging the dataset for these clusters and implementing hierarchical or soft-clustering techniques. A more sophisticated understanding of whether these clusters reflect stable subtypes within a wider visual spectrum or totally distinct stylistic categories might be possible with such enhancements.

Lastly, the conclusions' external validity is constrained by dataset bias. There is an issue with the 1,000-image dataset utilized in this study: it overrepresents Han and Qing Dynasty jade products. However, its representativeness is inadequate for some regional traditions, such the relatively unique Liangzhu Culture. The generalization of the research findings can be improved by expanding the dataset to cover those periods that have not been thoroughly examined. The accuracy and thoroughness of the clustering results can be increased with a more representative and varied dataset.

4.3. METHODOLOGICAL CONTRIBUTIONS

Three approaches—quantitative subjectivity, scale analysis, and interdisciplinary integration—are used in this study to demonstrate the methodological and applied value of computer vision in art history research.

First, it converts qualitative concepts like "symmetry" and "naturalism" into quantifiable attributes like motif density and line curvature. By eliminating the subjectivity present in conventional qualitative analysis, this quantitative method provides a more accurate and objective means of characterizing and analyzing the features of jade sculptures. Second, the system can find patterns and aggregate styles in big collections by using an autonomous detection and clustering method. Manual classification is often unlikely for large-scale cultural relic collections because of the project's enormous scale. Computer vision technology, on the other hand, is distinct. It is capable of processing and analyzing a lot of photos efficiently. It can highlight trends that human researchers would find difficult to identify. Third, by transparently verifying the expertise of annotators, it links AI with art history (Li, Gao, Yang). This method preserves the delicate cultural distinctions while guaranteeing the technology's rigor. This study preserved the integrity of cultural

interpretation while utilizing artificial intelligence for data analysis by incorporating art history specialists in the verification process.

5. CONCLUSION

5.1. SUMMARY OF FINDINGS

This study developed an analytical framework that integrates Jade-YOLOv9, quantitative feature extraction, and K-means clustering to examine stylistic variation in Chinese jade carving. Based on 1,000 high-resolution images, the proposed framework identified five major style categories—Normative Geometric, Free-Form Geometric, Natural Bionic, Minimalist Line, and Symbolic Pattern—with strong internal validity and high agreement with expert evaluation. The clustering results achieved an overall accuracy of 92.7%, demonstrating that computer vision can consistently recognize a wide range of artistic patterns that were previously described mainly through qualitative interpretation.

By transforming traditionally subjective visual concepts, such as symmetry, contour regularity, line curvature, and motif density, into measurable computational features, this study broadens and objectifies the conventional classification of jade carving styles. In this sense, the former descriptive understanding of jade carving traditions—ranging from the geometric precision of ritual jades to the concise expression of Han “Ba Dao” carving—can now be supported by repeatable quantitative evidence. Overall, the findings confirm the value of YOLOv9-based detection and K-means clustering for large-scale, data-driven research in traditional art and provide a reproducible methodological pathway for digital cultural heritage studies.

5.2. IMPLICATIONS FOR ART HISTORY AND CONSERVATION

This work gives insights for art historical interpretation and cultural relic conservation techniques in addition to new analytical methods for the study of the evolution of jade carving styles.

The classification framework presented in this paper aids scholars in re-understanding the distinctions and relationships between various style types from a quantitative standpoint at the level of art historical research. For instance, the formal features of ritual jade artifacts are empirically supported by the normative geometric group's high symmetry and high contour regularity; the identification of the minimalist line group and the symbolic pattern group offers a more objective analytical foundation for comparative studies of particular stylistic traditions. As a result, jade carving styles may now be examined in more detail within a framework that combines visual traits and historical context rather than depending just on empirical accounts.

The incorporation of quantitative elements at the level of cultural relic conservation opens up new avenues for digital archiving, identifying abnormal forms, and evaluating repair traces. The pertinent findings can serve as references for conservation staff, improving the foundation and targeting of conservation decisions, particularly when artifacts exhibit notable variations from their corresponding style categories in terms of symmetry, pattern density, or contour characteristics.

The visualization and quantitative analysis techniques developed in this paper can help advance the study of jade carving art from traditional textual interpretation to a comprehensive expression that combines images, data, and interpretation at the level of knowledge dissemination and teaching. They can also serve as methodological references for art history education and digital humanities research.

5.3. FUTURE DIRECTIONS

There is yet an opportunity to expand the use of artificial intelligence (AI) in art history research, even if the framework presented in this paper has shown good results in recognizing and clustering jade carving styles.

First, by adding details like age, locality, excavation backdrop, and cultural meanings, future studies could expand on the visual characteristics already in place. By improving the model's comprehension of intricate stylistic settings, this would lessen the errors that result from depending only on visual cues. This would facilitate the advancement of AI models from "image recognition" to "context-assisted interpretation."

Second, in order to better identify transitional relationships and hybrid forms between various styles, future research could investigate combining hierarchical clustering, soft clustering, or multimodal learning techniques to

address the common phenomena of stylistic overlap and blurred boundaries in contemporary jade carvings. By enhancing the model's ability to express complex artistic features, AI holds promise for providing more detailed technical support for the study of jade carving style genealogy.

Future studies can also investigate the possible applications of artificial intelligence (AI) in anomaly identification, auxiliary judgment in restoration, and digital archive organization in the field of cultural relic conservation and digital administration. For instance, pertinent models can serve as supplementary references for conservationists when artifacts substantially depart from established stylistic patterns in terms of symmetry, pattern density, or outline structure. This enhances the scientific rigor and effectiveness of cultural relic conservation efforts. Furthermore, future study can investigate the application potential of AI in art-assisted evaluation with the development of multimodal learning and explainable AI; nevertheless, pertinent judgments still need to be validated in conjunction with expert knowledge and particular historical context.

Lastly, using larger-scale and more varied data, future study can investigate the application of AI techniques in cross-period, cross-regional, and even cross-cultural comparative art studies. To guarantee that technical analysis always supports humanities research questions rather than being detached from particular cultural contexts, this approach still needs to be based on adequate data collection, trustworthy expert participation, and a clear art historical theoretical framework.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Ali, M. L., & Zhang, Z. (2024). The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection. Computers, 13(12), 336. https://doi.org/10.3390/computers13120336

Fang, A. (2024, December 31). Carving out a symbol of virtue: Modern jade craftsmanship. People's Daily Online. https://en.people.cn/n3/2024/1231/c90000-20260543.html

Florin, G. (2024). An analysis of research trends using artificial intelligence in cultural heritage. Electronics, 13(18), 3738. https://doi.org/10.3390/electronics13183738

Gao, C., Zhang, Q., Tan, Z., Zhao, G., Gao, S., Kim, E., & Shen, T. (2024). Applying optimized YOLOv9 for heritage conservation: Enhanced object detection in Jiangnan traditional private gardens. Heritage Science, 12(1), 31. https://doi.org/10.1186/s40494-024-01144-1

Hu, M. K. (1962). Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, 8(2), 179–187. https://doi.org/10.1109/TIT.1962.1057692

Hu, W., & Pei, C. (2025). Digital presentation of intangible cultural heritage: A case study of Beijing jade carving. Frontiers in Art Research, 7(3), 11–17. https://doi.org/10.25236/FAR.2025.070303

Lam, H. L. E. (2019). Representation of heaven and beyond: The Bi disc imagery in the Han burial context. Asian Studies, 7(2), 115–151. https://doi.org/10.4312/as.2019.7.2.115-151

Li, J., Yao, L., Hendriks, E., & Wang, J. Z. (2011). Rhythmic brushstrokes distinguish van Gogh from his contemporaries: Findings via automated brushstroke extraction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(6), 1159–1176. https://doi.org/10.1109/TPAMI.2011.203

Liang, M., & Tan, C. (2024). Jade and algorithm: Toward digital taxonomy of Chinese ornaments. Digital Applications in Archaeology and Cultural Heritage, 32, e00351. https://doi.org/10.1016/j.daach.2024.e00351

Lim, W. M. (2024). A typology of validity: Content, face, convergent, discriminant, nomological and predictive validity. Journal of Trade Science, 12(3), 155–179. https://doi.org/10.1108/JTS-03-2024-0016

Liu, R., Zhang, H., & Wang, L. (2022). A survey on deep learning in art and cultural heritage. ACM Computing Surveys, 55(9), 1–35. https://doi.org/10.1145/3510426

Lu, Y., & Liu, F. (2023). Evaluating motif density in digital jade carvings. Heritage Science, 11(1), 27. https://doi.org/10.1186/s40494-023-00811-w

Ma, H., & Theppituck, T. (2025). Classification of Chinese traditional jade carvings. International Journal of Sociologies and Anthropologies Science Reviews, 5(6), 657–666. https://doi.org/10.60027/ijsasr.2025.7579

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sun, S., Dustdar, S., Ranjan, R., Morgan, G., Dong, Y., & Wang, L. (2022). Remote sensing image interpretation with semantic graph-based methods: A survey. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 4544–4558. https://doi.org/10.1109/JSTARS.2022.3176612

Terven, J., Córdova-Esparza, D. M., & Romero-González, J. (2023). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv9 and YOLO-NAS. Machine Learning and Knowledge Extraction, 5(4), 1680–1716. https://doi.org/10.3390/make5040083

Wang, M., & Shi, G. (2020). The evolution of Chinese jade carving craftsmanship. Gems & Gemology, 56(1), 30–53. https://doi.org/10.5741/GEMS.56.1.30

Westlake, N., Cai, H., & Hall, P. (2016). Detecting people in artwork with CNNs. In European conference on computer vision (pp. 825–841). Springer International Publishing. https://doi.org/10.1007/978-3-319-46604-0_57

Wu, X., Yuan, Q., Qu, P., et al. (2025). Image-driven batik product knowledge graph construction. NPJ Heritage Science, 13, 20. https://doi.org/10.1038/s40494-025-01823-4

Xie, H., Liu, X., Chen, Y., & Sun, F. (2021). Combining CNNs and handcrafted features for ancient character recognition. Information Sciences, 562, 437–449. https://doi.org/10.1016/j.ins.2021.02.059

Yao, S., & Wu, Q. (2023). Neural-based generation of jade motif patterns. Pattern Recognition Letters, 169, 150–159. https://doi.org/10.1016/j.patrec.2023.03.007

Yu, T., Lin, C., Zhang, S., et al. (2022). Artificial intelligence for Dunhuang cultural heritage protection: The project and the dataset. International Journal of Computer Vision, 130, 2646–2673. https://doi.org/10.1007/s11263-022-01665-x

Zafeiropoulos, C., Tzortzis, I. N., Rallis, I., et al. (2021). Evaluating unsupervised clustering in cultural heritage monitoring. Journal on Computing and Cultural Heritage, 14(4), 1–17. https://doi.org/10.1145/3469006

Zhan, J., Meng, Y., Zhang, L., Li, K., & Yan, F. (2025). Research on computer vision in intelligent damage monitoring of heritage conservation: The case of Yungang cave paintings. NPJ Heritage Science, 13, 45. https://doi.org/10.1038/s40494-025-01945-9

Zhang, K., Liu, J., & Ma, Y. (2022). Research on clustering techniques for cultural object classification. Pattern Analysis and Applications, 25(3), 639–654. https://doi.org/10.1007/s10044-022-01011-w

Zhao, L., & Liu, T. (2023). Style transfer and preservation of Chinese jade carvings using neural networks. Journal of Cultural Heritage, 62, 195–204. https://doi.org/10.1016/j.culher.2022.10.005

Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(07), 12993–13000. https://doi.org/10.1609/aaai.v34i07.6999