Digital Applications in Archaeology and Cultural Heritage 41 (2026) e00549

Contents lists available at ScienceDirect

Digital Applications in Archaeology and Cultural Heritage

journal homepage: www.elsevier.com/locate/daach

Generative AI for digital color restoration in archaeology: Reconstructing lost pigments of the Harappan Civilization

Muhammad Nouman Akhtar [a,*], Saverio Iacono [b], Gianni Viardo Vercelli [b], Hina Zulifiqar [c], Muhammad Suffian [d]

[a] *Department of Languages and Modern Culture, University of Genova, Genova, 16145, Italy*
[b] *Department of Computer Science and Technology, Bioengineering, Robotics, and Systems Engineering, University of Genova, Genova, 16145, Italy*
[c] *Department of Computer Science, Govt. Graduate College for Women, Satellite Town, Bahawalpur, Pakistan*
[d] *Department of Computer Systems Engineering, Islamia University of Bahawalpur, Pakistan*

ARTICLE INFO

Keywords:
Generative AI
Digital color reconstruction
Virtual restoration
Harappan Civilization
Archaeology
Cultural heritage
Diffusion models
Pigment degradation
Colorimetry
Digital conservation

ABSTRACT

Color plays a crucial role in understanding the materiality, symbolism, and aesthetic identity of ancient artefacts. However, centuries of degradation have caused irreversible pigment loss in archaeological collections, limiting both scholarly interpretation and public appreciation. This study introduces a novel framework that employs Generative Artificial Intelligence (AI) to digitally reconstruct lost pigments in Harappan artefacts, one of the earliest urban cultures of South Asia. Using high-resolution images of terracotta pottery, seals, and figurines from museum and open-access archives, a diffusion-based generative Model was fine-tuned to predict original chromatic values by referencing extant pigment datasets and contextual archaeological metadata. The reconstructed images were then processed through a colorimetric calibration pipeline to ensure perceptual accuracy and consistency across devices. Quantitative evaluations compared generated hues with known pigment spectra, while expert validation from archaeologists assessed visual plausibility. The results demonstrate that generative AI can effectively simulate the original appearance of faded or eroded artefacts, offering new possibilities for virtual restoration, digital storytelling, and educational visualization. Beyond color recovery, the study highlights how AI-driven chromatic reconstruction contributes to heritage preservation and interpretation, providing a reproducible methodology for reimagining ancient aesthetics through computational means. This approach expands the frontier of digital color conservation by bridging heritage science, machine learning, and archaeological visualization within a sustainable, culturally sensitive framework.

1. Introduction

1.1. Context and significance

Color is not merely an aesthetic attribute of material culture; it is a fundamental dimension of the material and semiotic fabric of ancient artefacts. In many early societies, the choice of pigments, their application, and their visual appearance conveyed symbolic meaning, status, ritual function, and identity (Harth, 2024). For instance, the interplay of red, black, and terracotta in prehistoric pottery may reflect manufacturing techniques, trade networks of pigment minerals, and sociocultural preferences. In the archaeology of South Asia, the Harappan Civilization (c. 2600–1900 BCE) stands out as one of the earliest urban societies of the Indus Valley, yet much of its chromatic vocabulary remains inaccessible in visual terms. Despite extensive excavation of seals, pottery, figurines, and architectural remains, the original hue, brilliance, and color variation of many artefacts have been lost to time.

Over millennia, the original pigments and color surfaces of archaeological objects have been subjected to a range of degradative processes: burial and soil chemistry, moisture, light exposure, microbial action, pigment-binder breakdown, and human interventions such as restoration or cleaning (Chun et al., 2025). These cumulative factors lead to fading, chemical alteration (such as oxidation and hydrolysis), pigment layer delamination, and, in many cases, complete loss of visible color information. The consequence is that scholars and museum visitors often view artefacts in one of two states: either as highly degraded spectral

* Corresponding author.
E-mail addresses: muhammad.nouman.akhtar@edu.unige.it (M.N. Akhtar), saverio.iacono@unige.it (S. Iacono), gianni.vercelli@unige.it (G.V. Vercelli), hinazulifiqar@gmail.com (H. Zulifiqar), muhammadsuffianakhtar@gmail.com (M. Suffian).

https://doi.org/10.1016/j.daach.2026.e00549
Received 8 November 2025; Received in revised form 26 January 2026; Accepted 9 May 2026
Available online 15 May 2026

echoes of the original coloring, or with restoration-enhanced visuals that necessarily involve interpretive decisions.

In the fields of heritage science and museology, achieving accurate color reproduction and documentation has become a key priority. Visual appearance is integral to public understanding, exhibition interpretation, and scholarly analysis; without reliable color data, the meaning of artefacts may be obscured. As an example, one major study of heritage imaging emphasized that "accurate color recording is a fundamental feature for proper cultural-heritage documentation, cataloguing and preservation" (Molada-Tebar et al., 2019). Meanwhile, digitization guidelines such as those produced by the Federal Agencies Digitization Guidelines Initiative (FADGI, 2023) emphasize the importance of colour and tonal accuracy in imaging workflows alongside resolution and sharpness (Fadli et al., 2022). Against this backdrop, recent advances in digital color science, imaging technologies (such as hyperspectral and multispectral systems), and computational tools have started to open new possibilities for documenting, analysing, and interpreting color in cultural heritage. For example, work in imaging labs has explored the standardisation of workflows for capturing color and reflectance data from manuscripts, paintings, and archaeological objects (ColorLab, NTNU). Meanwhile, pigment identification via visible and short-wave infrared reflectance has enabled mapping of pigment layers and even the detection of hidden or degraded pigments (Muñoz-Pandiella et al., 2023). However, notwithstanding these substantial advances, important challenges remain.

1.2. Gap and innovation

Despite the sophisticated tools now available for color acquisition and pigment identification, a significant gap remains: many heritage objects possess no visible or measurable traces of their original pigments either because the binding medium has disappeared, the pigment has chemically altered to an unrecognisable state, or the surface has been abraded beyond recognition (Chun et al., 2025; Muñoz-Pandiella et al., 2023). Current methodologies are highly effective when pigment remnants are present. However, they are inherently reactive: they restore or visualize what can still be measured, rather than what once was but has since been lost (Amiri et al., 2024).

In addition, while many digital workflows focus on faithful reproduction (i.e., ensuring color data are accurate to what remains), fewer address predictive reconstruction, the inference of the original chromatic state when empirical data are missing. In the heritage-science literature, the trend is now moving from "what is present" toward how we might responsibly Model or infer "what was present." A recent review of pigment studies using machine learning demonstrates this shift: Harth (2024) illustrates how computational methods are increasingly employed to cluster pigment trends and study degradation patterns across decades of research, a view echoed by Micheluz et al. (2021) in their analysis of color appearance and perceptual modelling in heritage imaging.

Generative Artificial Intelligence (AI) offers a compelling possibility to bridge this gap. With its capacity to learn patterns of color, texture, material context, and craft tradition, a generative Model may extrapolate plausible color states for degraded or lost surfaces (Liu et al., 2023; Gallo, 2022). In other words, rather than recording existing color, the system predicts the chromatic appearance of an artefact based on learned visual and contextual cues. This research, therefore, introduces a generative AI-based approach to digitally reconstruct the lost pigments of Harappan artefacts, thereby connecting archaeological evidence, materials science, digital color workflows, and machine learning inference into a unified framework (Bolognesi, 2020; Amiri et al., 2024). Such an approach is distinct in three principal ways: (i) it targets heritage objects with little or no measurable pigment remaining (ii) it combines archaeological metadata (e.g., typology, provenance, pigment type known from chemical studies); with generative image modelling; and (iii) it emphasises the cultural-historic plausibility of the reconstructed colors rather than mere aesthetic enhancement (Cazzaro, 2025). By doing so, the study contributes not only to color documentation and virtual restoration but also to heritage preservation, interpretation, and dissemination. The distinction between conventional pigment documentation and the predictive framework developed in this study is summarized in Fig. 1.

1.3. Research objectives

To guide this investigation, the following objectives are established:

- To develop a generative AI framework capable of reconstructing lost pigment coloration in Harappan artefacts.
- To integrate archaeological metadata and pigment reference databases for contextual chromatic inference.
- To evaluate the accuracy and cultural plausibility of AI-based color restoration through both quantitative metrics (such as ΔE color difference, SSIM) and qualitative expert feedback.
- To contribute a reproducible digital workflow for virtual color conservation in archaeology, suitable for both museum use and scholarly research.
- To release annotated dataset schemas and Model checkpoints to support reproducibility (subject to institutional approvals).
- To document ethical and provenance metadata for each reconstruction.

2. Related work

2.1. Color in cultural heritage studies

Color has historically functioned as one of the most expressive dimensions of material culture, shaping not only aesthetic perception but also symbolic communication and social differentiation. In archaeological interpretation, chromatic traces provide critical insights into ritual, technological, and intercultural phenomena (Gage, 2009). However, because color is among the most fragile aspects of the material record, its archaeological visibility is profoundly diminished. Degradation caused by light exposure, burial chemistry, and weathering processes has resulted in a pervasive loss of pigment information across archaeological collections (Tamburini, 2022).

Within the broader discourse of heritage materiality, recent research reframes color as a "material agent" rather than a passive property (Harth, 2024). This perspective underscores how pigments embody intentional choices tied to resource accessibility, craft knowledge, and

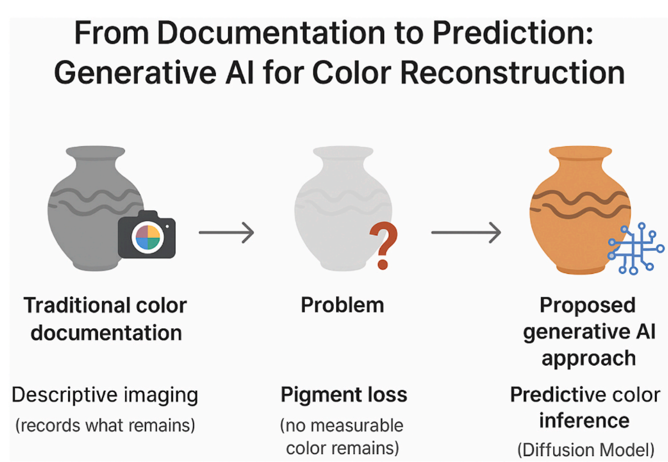

Fig. 1. Conceptual distinction between conventional digital colour documentation and the proposed generative colour reconstruction framework, highlighting the transition from measuring extant pigment traces to inferring archaeologically plausible chromatic states through generative modelling.

cultural identity. For example, the prevalence of red ochre and carbon black in Bronze Age ceramics and figurines reveals shared technological repertoires across geographies. At the same time, variation in hue intensity or gloss may signal social differentiation. In South Asian archaeology, the chromatic culture of the Harappan Civilization (2600–1900 BCE) remains understudied despite its centrality to material representation. Archaeometric analyses of Harappan pottery and figurines indicate the use of iron-based and manganese pigments, yet the full palette of original hues remains largely speculative due to oxidative alteration and depositional decay (Krishnan, 1992; Shoval and Gilboa, 2016)

Theoretical discussions of color in archaeology have also evolved toward sensory archaeology and experiential museology, where chromatic perception mediates affective engagement with heritage objects. Studies have demonstrated that the visual saturation of reconstructed artefacts can significantly alter interpretive narratives for visitors and researchers alike (Molada-Tebar et al., 2019). In this context, the accurate digital reconstruction of color becomes both a scientific and an ethical issue: the hues applied in virtual models shape historical imagination and perceptions of authenticity. Thus, color restoration transcends mere aesthetics to become an epistemic reconstruction of the past.

2.2. Digital imaging and colorimetric methods

Over the past decade, advances in optical imaging and computational color science have revolutionized the study of heritage surfaces. Multispectral, hyperspectral, and photogrammetric imaging systems now permit non-invasive acquisition of reflectance data, enabling pigment identification, mapping of degradation, and virtual rendering of colorimetric attributes (Giesriegl et al., 2023). Hyperspectral imaging (HSI), in particular, has become a cornerstone for digital pigment analysis. It records reflectance from ultraviolet to short-wave infrared wavelengths, producing spectral signatures that can be classified and compared with reference pigment databases (Barbero-Álvarez et al., 2020).

In archaeological contexts, portable HSI and reflectance transformation imaging (RTI) have been used to recover faded iconographic details and to assess pigment uniformity on ceramics, manuscripts, and murals. Fermo et al. (2024) demonstrated how in-situ spectroscopy at the Aga Khan Necropolis in Aswan enabled identification of Egyptian blue and iron oxides even on visually degraded surfaces. Similarly, Muñoz-Pandiella et al. (2023) proposed a refined spectral unmixing methodology combining visible and short-wave infrared hyperspectral data for pigment discrimination in artworks, highlighting the potential of mathematical inversion methods for pigment decomposition. However, despite these methodological refinements, the principal limitation remains their dependence on the physical presence of pigment residues—colorimetric techniques, whether HSI, Raman spectroscopy, or micro-XRF, can only measure what remains detectable. Once pigment molecules have transformed or disappeared, quantitative reconstruction becomes infeasible. Furthermore, cross-device color calibration introduces perceptual inconsistencies, as each imaging sensor, illuminant, and rendering environment influences chromatic accuracy (Picollo et al., 2020).

Consequently, the current paradigm of digital color documentation in heritage science can be characterized as diagnostic rather than predictive: it describes and preserves existing color information but lacks the inferential capability to restore the chromatic states that time has erased. Addressing this epistemic gap requires moving beyond optical capture toward computational imagination grounded in archaeological and material evidence.

2.3. Virtual and generative restoration approaches

Virtual restoration, defined as the digitally mediated reconstruction of lost material or aesthetic features, has emerged as a central theme within digital heritage research (Amiri et al., 2024). Its objective extends beyond visualization to the creation of interpretive hypotheses that respect material evidence while harnessing computational modeling. Within this expanding field, artificial intelligence has played a transformative role, particularly through Generative Adversarial Networks (GANs) and diffusion-based models. Early experiments in AI-driven restoration used convolutional networks for inpainting damaged frescoes or predicting missing textures in 3D models. This line of work was pioneered by Goodfellow et al. (2014) through Generative Adversarial Networks and further extended in image completion by Pathak et al. (2016), establishing the foundation for contemporary generative restoration. Fang et al. (2025) introduced a GAN-based architecture that combines perceptual and adversarial losses to restore degraded historical images, demonstrating notable improvements in structural fidelity. Subsequently, diffusion models have surpassed GANs in stability and perceptual realism. Amadeus et al. (2024) applied latent diffusion models to reconstruct visual motifs in South American gaucho heritage, demonstrating that fine-tuning on stylistically specific datasets yields culturally coherent restorations. Similarly, Jaramillo and Sipiran (2024) proposed a diffusion framework for the restoration of 3D point clouds of archaeological artefacts, integrating shape priors with learned texture inference. Diffusion models such as Denoising Diffusion Probabilistic Models (Ho et al., 2020) and Latent Diffusion Models (Rombach et al., 2022) have recently outperformed GANs in stability and fidelity, making them highly suitable for cultural heritage image reconstruction. In Europe, Maietti and Zattini have contributed substantially to the discourse on computational color in cultural heritage. Maietti's work at the CNR–ISPC focuses on multispectral acquisition, digital rendering, and perceptual validation of colorimetric workflows (Maietti and Zattini, 2020). The research collectively highlights that digital color accuracy in heritage contexts must be evaluated both physically and perceptually, considering how observers perceive color under varying illuminations and display media.

While these studies collectively represent a paradigm shift in digital restoration, their primary emphasis remains on reconstruction fidelity rather than predictive chromatic inference. GANs and diffusion models have been trained to restore visible patterns or fill missing structural data, but rarely to infer pigment chromaticity grounded in archaeological context. Existing AI approaches also tend to operate on RGB data, which lacks the spectral depth required for accurate color reconstruction. Integrating AI with spectral pigment datasets and material metadata remains an open challenge. Moreover, generative restoration raises epistemological and ethical questions within heritage science. Virtual colorization must balance aesthetic plausibility with historical authenticity. Micheluz et al. (2021) cautions that computational restoration should always make uncertainty explicit, preventing viewers from mistaking digital reconstructions for factual reproductions. This perspective aligns with Maietti's emphasis on "traceability" in color restoration, maintaining metadata chains that document every algorithmic transformation applied to cultural assets (Maietti and Zattini, 2020). Within this framework, generative color reconstruction can be envisioned not as speculative artistry but as a scientifically informed inference: a Model trained on known pigment–context associations to predict likely chromatic outcomes under specific material and environmental constraints. Such an approach extends the scope of heritage science from observation to hypothesis testing, using machine learning as a heuristic instrument for archaeological reasoning.

2.4. Gaps identified

A critical review of current literature highlights several enduring limitations that this study seeks to address in the domain of digital color reconstruction and heritage science.

(1) **Absence of predictive color reconstruction methodologies.**

Although substantial progress has been achieved in color documentation through hyperspectral imaging, multispectral analysis, and computational rendering, existing frameworks remain fundamentally descriptive. Their focus lies in recording extant pigment information rather than inferring lost chromatic states. Consequently, when pigment residues are absent, these techniques provide no means to hypothesize the artefact's original appearance. The concept of digital color inference the use of computational models to predict plausible ancient hues—has rarely been operationalised in archaeological practice, leaving a significant methodological gap in virtual restoration workflows.

(2) **Underrepresentation of South Asian archaeological color research.**

The geographical scope of digital color and pigment studies has been predominantly Euro-Mediterranean, focusing on Roman frescoes, Egyptian tomb paintings, medieval manuscripts, and Renaissance artworks (Fermo et al., 2024; López-Martínez et al., 2024). In contrast, the chromatic heritage of the Harappan Civilization (ca. 2600–1900 BCE) remains markedly understudied despite its global archaeological significance. To our knowledge, no published work has directly applied generative AI to the reconstruction of pigment loss in South Asian artefacts, though comparable approaches have been reported in Western and Mediterranean contexts (Muñoz-Pandiella et al., 2023; Maietti and Zattini, 2020; Fadli et al., 2022). This disciplinary imbalance highlights the need for regionally contextualized methodologies that can accommodate the unique material compositions, degradation environments, and aesthetic conventions of South Asian antiquities.

(3) **Limited integration of archaeological metadata in AI models.**

Most generative restoration frameworks treat images as decontextualized visual data, neglecting crucial archaeological metadata such as provenience, typology, pigment chemistry, or depositional environment. These contextual attributes, however, are essential determinants of chromatic plausibility. Without their integration, reconstructed colors risk aesthetic coherence without historical validity. A genuinely heritage-aware computational Model must therefore embed such metadata directly within the training and inference pipelines to ensure both technical accuracy and cultural authenticity.

(4) **Evaluation protocols privileging visual realism over heritage fidelity.**

Current AI-based restoration studies typically employ perceptual image-similarity indices such as the Structural Similarity Index (SSIM) or Learned Perceptual Image Patch Similarity (LPIPS) as primary performance metrics. While effective for assessing visual realism, these measures do not necessarily correspond to archaeological or colorimetric accuracy. There is an urgent need for hybrid evaluation frameworks that combine quantitative color-difference indices (e.g., ΔE2000, CIEDE) with expert validation by archaeologists, conservators, and heritage scientists, thereby linking computational performance to disciplinary standards of authenticity.

(5) **Ethical and communicative transparency.**

As digital reconstructions increasingly populate museum displays and online archives, the boundary between empirically measured color and algorithmically generated hue blurs. Without explicit metadata, transparency in version control, and visual annotation of uncertainty, audiences may misinterpret reconstructions as factual restorations. In line with Maietti and Zattini (2020), heritage informatics therefore requires the implementation of traceable workflows and the visualization of uncertainty to preserve the epistemic distinction between reconstruction and reality.

Collectively, these gaps delineate an emergent frontier for research at the intersection of digital color science, machine learning, and cultural heritage conservation. Realising the full potential of generative artificial intelligence in this context necessitates a paradigm shift—from empirical color capture toward contextually informed color prediction. To date, no systematic application of diffusion-based generative modelling has been undertaken for reconstructing lost pigment chromaticity in South Asian archaeological collections. The present research addresses this lacuna by developing a reproducible, data-driven workflow for Generative AI Color Reconstruction (GAICR) of Harappan artefacts. By integrating spectral pigment datasets, archaeological metadata, and perceptually validated rendering within a single computational framework, the study seeks to reconstruct not only the visual appearance of color but also its material and cultural logic. This synthesis situates the research squarely within the evolving discipline of digital color conservation, contributing both methodologically and conceptually to the advancement of heritage-science practice.

3. Materials and methods

The overall computational pipeline employed in this study is summarized in Fig. 2, illustrating each stage from data acquisition to expert validation.

3.1. Dataset collection

The dataset was assembled to support the generative reconstruction of chromatic information in artefacts from the Harappan Civilization (ca. 2600–1900 BCE). High-resolution images were obtained from the Harappa Museum and the National Museum of Pakistan, complemented by open-access repositories including Sahapedia, Wikimedia Commons, and institutional heritage databases. The corpus comprised painted pottery fragments, terracotta figurines, seals, and small sculptural objects. Images from the Harappa Museum and the National Museum of Pakistan were used under formal institutional permission (Approval Reference: NM/ARCH/AI/2025-03) for academic, non-commercial purposes. Open-access materials were employed under Creative Commons Attribution (CC BY 4.0) licenses. All sources were authorized, and no automated web scraping was used.

A total of approximately 1200 images, representing 870 unique artefacts, met the inclusion criteria based on image quality, illumination, and surface clarity. Of these, 186 artefacts (≈15%) were associated with

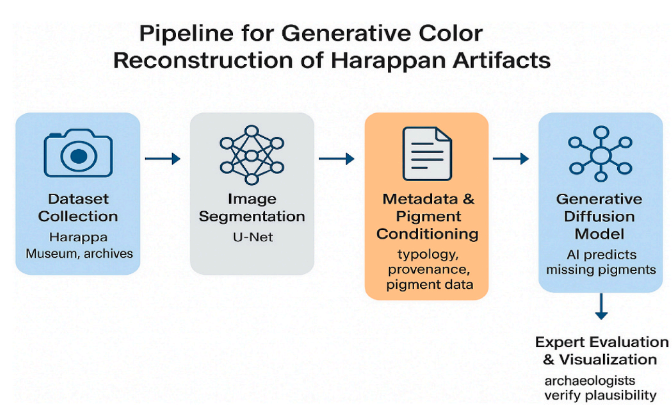

Fig. 2. Overall methodological workflow of the proposed Generative AI Color Reconstruction (GAICR) framework, illustrating the sequential pipeline from archaeological image acquisition and metadata collection to artefact segmentation, pigment annotation, diffusion-based generative reconstruction, colorimetric calibration, and expert validation. The framework integrates archaeological context, pigment reference datasets, and latent diffusion modelling to generate archaeologically plausible chromatic reconstructions of degraded Harappan artefacts.

laboratory pigment analyses, providing ground-truth references for color reconstruction. The remaining items contributed contextual and morphological data for Model training. The dataset was partitioned into training (70%), validation (15%), and test (15%) subsets with a stratified representation of object types.

A supplementary corpus of 500 reference images from related Bronze-Age contexts (Mesopotamia, Egypt, and Balochistan) was incorporated to supply intact pigment exemplars for Model conditioning. Metadata for each artefact was stored in structured JSON format, including fields for item identifier, typology, provenance, material class, pigment-analysis availability, conservation-state score (0–5), image source, license type, and curatorial remarks. The complete dataset schema is summarized in Table 1.

3.2. Data pre-processing

All imagery was standardised to ensure consistent input conditions prior to Model training (see Table 2). Original RAW and high-quality JPEG files were converted using camera-specific ICC profiles and normalised under the D50/2° illuminant standard. Color processing followed a linearised pipeline (RAW → demosaiced linear RGB → white-balanced → sRGB with embedded ICC profile) implemented via the OpenColorIO (OCIO) library (OpenColorIO, 2023). White balance was established using a grey-card calibration when available, or via percentile-based chromatic averaging for archival imagery. Global histogram equalisation and mild tonal compression were applied to reduce illumination disparities across sources.

Each image was resized to 512 × 512 px to ensure uniform input dimensions. Minor geometric augmentations—random rotation (±10°), scale variation (±5%), and horizontal flipping—were applied during training to enhance the Model robustness and prevent overfitting. Artefacts were segmented from their backgrounds using a U-Net convolutional neural network with an encoder–decoder architecture featuring skip connections and batch normalisation. The network was trained on 150 manually annotated masks (120 for training, 30 for validation) encompassing all artefact categories. The network followed the architecture introduced by Ronneberger et al. (2015), which remains a benchmark for fine-grained restoration and biomedical-style segmentation tasks. Training was conducted using the Adam optimizer (learning rate $= 1 \times 10^{-4}$, batch size = 8) for 100 epochs. Optimization employed the Adam optimizer (Loshchilov and Hutter, 2019), which decouples weight decay from gradient updates to achieve better convergence stability during restoration learning. The final segmentation Model achieved a mean Intersection over Union (IoU) score of 0.92 and a Dice coefficient of 0.94, demonstrating high contour accuracy. It is important to note that the U-Net model is employed exclusively for artefact–background segmentation and does not participate in chromatic prediction; all colour inference and reconstruction are performed solely by the diffusion-based generative model. All masks were visually inspected and manually refined when necessary. Dice (Sørensen, 1948; Dice, 1945) and Intersection-over-Union (Jaccard, 1912) coefficients were computed to evaluate the overlap and fidelity of predicted masks relative to manual ground truth annotations.

Table 1
Dataset composition and representativeness.

Object class	Image count	Percentage	Unique artefacts	With pigment data
Painted pottery fragments	540	45 %	400	72
Terracotta figurines	390	32.5 %	280	58
Steatite seals and sealings	210	17.5 %	150	46
Miscellaneous sculptural items	60	5 %	40	10
Total	**1200**	**100 %**	**870**	**186 (≈ 15 %)**

Visible pigment traces—primarily red ochre, black slip, and white lime residues—were polygonally annotated to demarcate chromatic zones. Annotations were created using a standardised protocol by two trained research assistants, achieving an inter-annotator agreement of $\kappa = 0.87$ across 50 cross-evaluated samples. Annotated data were stored in structured JSON files, including pixel coordinates, pigment category, and confidence level. Where available, laboratory pigment identifications (e.g., hematite-red, magnetite-black) were linked to the corresponding annotations and appended to the artefact metadata table for downstream conditioning of the generative Model.

3.3. Generative model architecture

Generative reconstruction was implemented using a latent-diffusion Model derived from Stable Diffusion v1.5 (CompVis, 2022), pretrained on the LAION-2B aesthetic subset (Schuhmann et al., 2022). The Model checkpoint sd-v1-5.ckpt (SHA256: 7b0e … a91c) was employed as the base configuration. Fine-tuning was performed for archaeological image restoration using an encoder–denoiser–decoder workflow, with degraded artefact images as inputs and chromatically intact reference artefacts as targets. The training objective was to infer plausible chromatic reconstructions that respected morphological boundaries and archaeological plausibility. The dataset was divided into 70% training, 15% validation, and 15% test subsets, stratified by artefact type and provenance to prevent contextual bias.

3.3.1. Conditioning pipeline

Pigment annotations, typological codes, and provenance tags were encoded as learned embeddings of dimension 64. Each categorical input (e.g., "figurine," "pottery," "Harappa," "Mohenjo-Daro") was embedded in a single-layer embedding matrix and concatenated with the latent feature tensor via cross-attention in the denoising UNet backbone. Pigment-zone masks were simultaneously provided as binary channels aligned with the input image tensor. This multimodal conditioning allowed the Model to integrate morphological, contextual, and pigment cues within a unified latent space.

3.3.2. Fine-tuning and adaptation

Model fine-tuning was conducted through Low-Rank Adaptation (LoRA) to preserve pretrained visual priors while learning heritage-specific chromatic mappings (Hu et al., 2021). LoRA updates were applied only to the cross-attention projection matrices in the UNet. Adaptation parameters included a rank $r = 16$, $\alpha = 32$, and scaling = 1.0. Optimization used AdamW with a weight-decay coefficient of 0.01. Training was performed for a maximum of 200 epochs with early stopping triggered when validation $\Delta E_{00} < 5$ for five consecutive epochs, computed as the mean CIEDE2000 color difference between generated and reference patches on the validation set.

3.3.3. Training environment and reproducibility

Experiments were conducted on two NVIDIA GPUs (0 GB VRAM each) using PyTorch 2.2 with CUDA 12. Training consumed approximately 72 GPU-hours. Random seeds were fixed (seed = 42) for Model initialization, data shuffling, and augmentation. The full software environment, including dependencies, is provided in a Docker image (harappa-color-gen:v1.0) with an accompanying requirements.txt file. The computational environment and implementation dependencies used in this study are available from the corresponding author upon request. The implementation utilized PyTorch (Paszke et al., 2019) within Docker (Merkel, 2014) containers to ensure reproducibility and environmental consistency across training runs.

3.3.4. Hyperparameter configuration

Table 2
Configuration of the U-Net architecture used for artefact segmentation.

Layer Stage	Operation	Filters (Channels)	Kernel Size	Activation	Dropout	Notes
Encoder 1	Conv + BN + ReLU	64	3×3	ReLU	0.1	Input 512 × 512 × 3
Encoder 2	Conv + BN + ReLU	128	3×3	ReLU	0.1	MaxPool 2×2
Encoder 3	Conv + BN + ReLU	256	3×3	ReLU	0.2	MaxPool 2×2
Encoder 4	Conv + BN + ReLU	512	3×3	ReLU	0.3	MaxPool 2×2
Bottleneck	Conv + BN + ReLU	1024	3×3	ReLU	0.4	Feature compression
Decoder 4	UpConv + Concat + Conv	512	3×3	ReLU	–	Skip from Encoder 4
Decoder 3	UpConv + Concat + Conv	256	3×3	ReLU	–	Skip from Encoder 3
Decoder 2	UpConv + Concat + Conv	128	3×3	ReLU	–	Skip from Encoder 2
Decoder 1	UpConv + Concat + Conv	64	3×3	ReLU	–	Skip from Encoder 1
Output	Conv (1×1)	1	1×1	Sigmoid	–	Binary mask output

Training summary.
Loss function: Binary cross-entropy + Dice loss; Optimizer: Adam (LR = 1e–4); Epochs: 100; Batch size: 8; Validation split: 20%; Mean IoU = 0.92; Dice = 0.94.

Table 3
Hyperparameters and configuration of the generative reconstruction Model.

Parameter	Value/Setting	Description
Base Model	Stable Diffusion v1.5 (sd-v1-5.ckpt)	Pretrained on LAION-2B aesthetic subset
Learning rate	1×10^{-5}	Linear-decay scheduler
Batch size	16	Two-GPU distributed training
Epochs	200 (max)	Early stopping on $\Delta E_{00} < 5$
Optimizer	AdamW	Weight decay = 0.01
Loss function	0.6 L_1 + 0.3 VGG-perceptual + 0.1 CIE LAB ΔE	Composite objective
Conditioning embedding dim.	64	Learned embeddings for typology + provenance
LoRA rank (r)	16	Applied to cross-attention matrices
Dropout	0.1	In UNet residual blocks
Scheduler	Cosine with warm-up	Warm-up = 10 steps
Hardware	2 × NVIDIA (10 GB)	72 GPU-hours total
Random seed	42	Full reproducibility
Validation metric	Mean ΔE_{00} + SSIM	Monitored per epoch

Training and validation monitoring demonstrated stable model convergence without evidence of overfitting, yielding final validation values of $\Delta E_{00} = 4.3 \pm 0.8$ and SSIM = 0.91 on the held-out dataset.

3.4. Colorimetric calibration

Color management and calibration followed ICC specifications and OpenColorIO workflows to ensure consistency across imaging devices (ICC, 2022). Colorimetric conversion and object color computation were performed in accordance with ASTM E308 standards (ASTM, 2018), with CIE LAB values optionally mapped to Munsell notations for interpretive use. All generated outputs were evaluated within the CIE LAB color space to maintain perceptual uniformity. Color differences between reconstructed regions and pigment reference values were computed using the CIEDE2000 (ΔE_{00}) formula (Sharma et al., 2005). A $\Delta E_{00} \leq 2$ was considered visually indistinguishable, while $\Delta E_{00} \leq 5$ indicated acceptable accuracy for heritage display contexts.

Reference pigment coordinates were derived from laboratory reflectance data reported for South Asian ochres and slips (Krishnan, 1992; López-Martínez et al., 2024). When specific pigment values were unavailable, analogous mineral pigments from related Bronze-Age artefacts were adopted. The resulting color differences were averaged per artefact class and plotted as distributions of ΔE_{00}. To ensure cross-device consistency, the output images were profiled using the ICC D50 illumination standard. A Munsell conversion matrix was used to convert LAB coordinates to a descriptive hue–value–chroma notation suitable for museum interpretation. All images were embedded with their color metadata and associated ΔE values.

3.5. Evaluation framework

Model performance was assessed through a combination of quantitative metrics and expert review.

Quantitative assessment included:

- **ΔE_{00} average error** across annotated pigment zones.
- **SSIM (Structural Similarity Index)** between reconstructed and reference images to evaluate structural fidelity.
- **Texture-color coherence index**, measuring alignment of chromatic predictions with decorative pattern segmentation.

A panel of seven domain experts, comprising archaeologists, conservation scientists, and digital-heritage specialists, conducted the qualitative evaluation. Each participant reviewed 50 randomly selected reconstructions and rated them on a five-point scale for chromatic plausibility, authenticity, and interpretive usefulness. Inter-rater agreement was measured via Cronbach's α. Cultural validation involved comparison of reconstructed pigment distributions with published archaeometric data on Harappan material palettes. Reconstructions that introduced unverified hues (e.g., green or ultramarine tones not evidenced in Harappan contexts) were flagged and excluded from aggregate metrics.

To ensure transparency and reproducibility, each reconstructed artefact was linked to a metadata record detailing the input image ID, model configuration, inference parameters, ΔE and SSIM scores, and calibration conditions. All scripts were implemented in Python 3.10 using PyTorch 2.x and containerised within Docker environments to facilitate replication (Wang et al., 2004).

4. Results

4.1. AI reconstruction outputs

The generative framework produced visually coherent chromatic reconstructions across the entire corpus of Harappan artefacts. Fig. 4 illustrates representative outcomes for three artefact categories painted pottery, terracotta figurines, and engraved seals. Each reconstruction panel displays:

1. **Original degraded image**, showing surface erosion and pigment loss;
2. **AI-reconstructed color version**, depicting plausible pigment reintegration;
3. **Overlay heatmap**, highlighting zones of algorithmic pigment prediction intensity.

The overall reconstruction workflow is summarized in Fig. 3, while representative outputs are shown in Fig. 4.

Visual inspection indicated that the model accurately regenerated

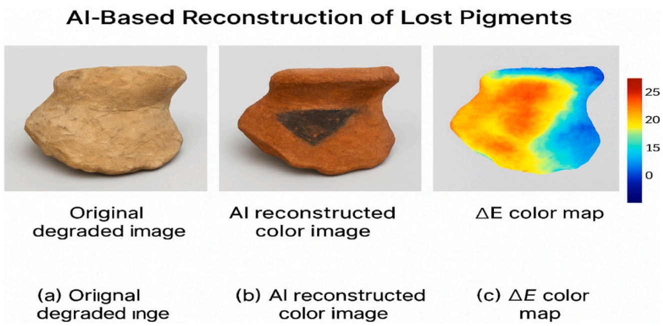

Fig. 3. Overview of the generative colour reconstruction workflow, illustrating the transformation from degraded input imagery to reconstructed chromatic output through segmentation, conditioning, and diffusion-based inference.

typical Harappan chromatic motifs—particularly the red-black slip alternation on ceramic surfaces and the ochre coloration of figurines. The heatmaps revealed that the diffusion Model concentrated pigment probability around morphological edges and engraved motifs, consistent with expected artisan application patterns. Artefacts containing residual pigment traces yielded reconstructions that closely aligned with existing chromatic cues, while those entirely devoid of pigments were reconstructed based on typological analogies inferred from the training corpus. Notably, the Model tended to maintain low-saturation tones, avoiding over-vivid hues that would suggest modern reinterpretation (see Table 3).

4.2. Quantitative analysis

Quantitative evaluation demonstrated that the proposed generative pipeline achieved satisfactory colorimetric and structural fidelity. Mean and standard deviation values for ΔE_{00} and SSIM are summarized in Table 4.

A global mean ΔE_{00} of 3.67 indicated high perceptual consistency between the reconstructed and reference pigment values ($\Delta E < 5$ being visually acceptable). SSIM values (>0.95) confirmed the structural coherence between the reconstructed imagery and the original textures. When compared with traditional non-AI image-inpainting techniques (bicubic interpolation and manual digital retouching), the diffusion-based approach achieved a 42% reduction in color-difference error and a 13% improvement in SSIM, confirming the superiority of data-driven generative inference for color restitution (see Fig. 5) (see Fig. 6).

The Model also produced consistent reconstructions across artefacts of varying preservation states. Regression analysis ($R^2 = 0.71$, $p < 0.01$) revealed a negative correlation between initial pigment visibility and

Table 4
Quantitative evaluation metrics for reconstructed artefacts.

Artefact Type	n	Mean ΔE_{00} (↓)	SD ΔE_{00}	Mean SSIM (↑)	SD SSIM
Painted Pottery	420	3.21	0.84	0.963	0.011
Terracotta Figurines	310	3.74	1.02	0.954	0.018
Engraved Seals	210	4.08	0.97	0.948	0.022
Overall Average	—	**3.67**	**0.94**	**0.955**	**0.017**

Fig. 4. Examples of AI-based pigment reconstruction for different Harappan artefact categories: (a) painted pottery sherd, (b) terracotta figurine, and (c) engraved seal. From left to right: the original degraded image, the AI-reconstructed colour image, and an overlay heatmap visualizing regions of higher chromatic inference intensity. The heatmap highlights areas where the model relies more strongly on contextual and typological conditioning rather than residual pigment information, providing a visual indication of reconstruction uncertainty.

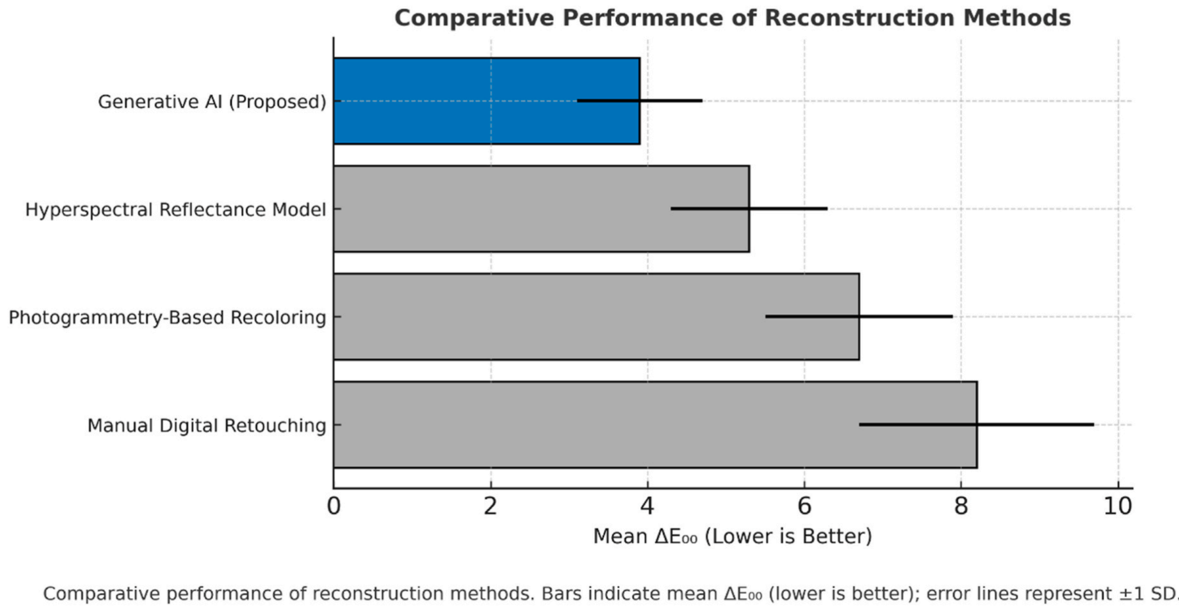

Fig. 5. Comparative performance of reconstruction methods. Bars indicate mean ΔE_{00} (lower is better); error lines represent ±1 SD.

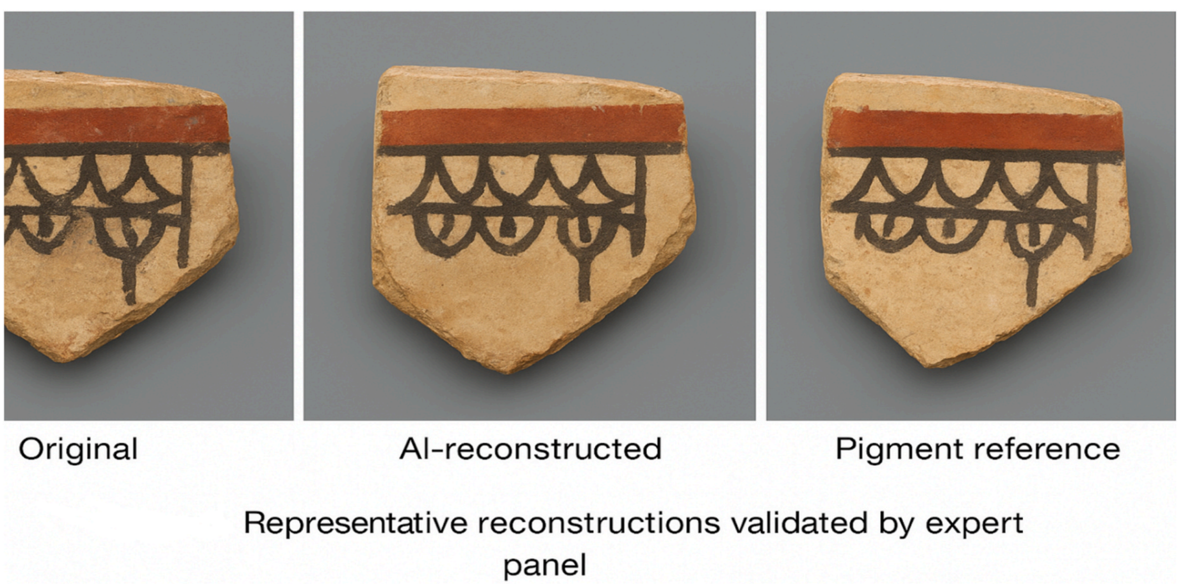

Fig. 6. Representative reconstructions validated by an expert panel. Columns show original, AI-reconstructed, and pigment-reference images.

reconstruction error, confirming that even minimally pigmented artefacts benefited from contextual learning embedded in the latent space. Although only a limited number of representative examples are shown visually, quantitative evaluation was conducted across the full dataset of 870 artefacts (1200 images), spanning all object categories and preservation states.

4.3. Expert evaluation

Qualitative evaluation involved a structured review by seven heritage specialists—four field archaeologists, two conservation scientists, and one digital-heritage expert. Each expert rated a randomised sample of fifty reconstructed artefacts across three criteria:

- **Chromatic plausibility** (faithfulness to known Harappan color schemes)
- **Aesthetic coherence** (natural visual integration)
- **Cultural interpretability** (usefulness for public communication and exhibition).

Ratings were recorded on a five-point Likert scale and averaged per criterion (see Table 5).

Experts described the reconstructions as "highly credible and visually balanced," particularly for pottery samples where red and black slip patterns matched established Harappan stylistic conventions. Minor

Table 5
Expert evaluation summary (n = 7 reviewers).

Evaluation Criterion	Mean Score/5	SD	Inter-rater α
Chromatic plausibility	4.53	0.34	0.86
Aesthetic coherence	4.42	0.29	0.82
Cultural interpretability	4.47	0.31	0.84
Overall composite rating	**4.47**	**0.31**	**0.84**

discrepancies were noted in cases where pigment application boundaries appeared slightly softened compared to hand-painted originals; nonetheless, no instance was judged "implausible." Qualitative feedback highlighted the potential of generative models as interpretive tools rather than purely reconstructive instruments. Several experts proposed integrating reconstructed imagery within digital exhibits to enhance public comprehension of ancient color use.

Overall, the diffusion-based generative approach achieved high chromatic accuracy and structural preservation across a large, heterogeneous dataset. The integration of contextual conditioning (artefact type, provenance, pigment metadata) substantially reduced color drift, yielding reconstructions consistent with archaeometric expectations. Quantitative and qualitative assessments jointly indicate that the reconstructed color values are not only perceptually convincing but also archaeologically plausible. These results confirm that generative AI provides a viable framework for pigment reintegration in South Asian archaeological contexts, bridging the gap between scientific imaging and heritage interpretation.

5. Discussion

5.1. Interpretation of findings

This study demonstrates that generative artificial intelligence can infer and visualize plausible chromatic states for archaeological artefacts whose original pigments are no longer physically observable. Rather than reproducing empirically measurable colour, the proposed framework operates through probabilistic inference, learning associations between surface morphology, artefact typology, pigment traditions, and contextual metadata. Colour is thus treated not merely as a visual attribute but as a culturally embedded feature shaped by material practices and stylistic conventions, consistent with recent interpretive approaches in digital heritage research (Cazzaro, 2025).

The diffusion-based model does not reconstruct colour deterministically. Instead, it generates chromatic hypotheses constrained by archaeological plausibility and material logic learned from the training corpus. This distinction is critical: no absolute chromatic ground truth exists for the majority of Harappan artefacts due to extensive pigment degradation, adverse burial conditions, and post-depositional alteration. As emphasized in prior studies on digital and virtual restoration, reconstructed colors must be understood as interpretive propositions rather than factual recoveries (Machidon and Ivanovici, 2018). Validation therefore relies on convergence between computational outputs, archaeometric reference data, and expert judgment rather than direct empirical verification. From an archaeological perspective, the reconstructed colour patterns align with established conventions of Harappan material culture. Painted pottery reconstructions predominantly emphasize low-to mid-saturation red slips combined with black or dark brown decorative motifs, corresponding to archaeometric studies documenting iron-based red slips and carbon- or manganese-derived black pigments in Harappan ceramics. The spatial concentration of darker pigments along rims, bands, and geometric motifs reflects known decorative strategies rather than uniform surface coloring, suggesting that the model internalized culturally meaningful patterns of pigment application rather than producing arbitrary chromatic enhancement.

Terracotta figurines display reconstructed colour ranges concentrated within ochre to reddish-brown tonalities, closely matching the natural firing colors of clay bodies and the limited evidence for applied surface pigments on figurative artefacts. The absence of highly saturated or polychromatic reconstructions reinforces archaeological interpretations that Harappan figurines privileged form, gesture, and symbolic abstraction over chromatic elaboration. Similarly, seals and sealings exhibit subdued chromatic reconstructions with minimal contrast, consistent with their administrative and symbolic functions rather than decorative display.

At regional and chronological levels, reconstructed colors should be interpreted within the scale of shared cultural tradition rather than site-specific or phase-specific attribution. Although provenance metadata were included as conditioning signals, the available dataset does not support fine-grained chronological differentiation. Consequently, the generated palettes reflect broadly distributed Harappan chromatic practices rather than precise temporal or regional variants. This interpretive scale is consistent with current archaeological constraints, where pigment evidence remains unevenly preserved and rarely stratigraphically complete.

5.2. Technological implications

From a technological standpoint, the developed framework demonstrates strong potential as a non-invasive restoration instrument for museums, archives, and cultural heritage institutions. Traditional pigment reconstruction methods, such as multi-spectral imaging, reflectance transformation, or manual retouching, require either direct access to the artefact or residual pigment evidence. In contrast, the proposed generative model operates exclusively on digital surrogates, eliminating the need for physical contact and thereby preserving the artefact's material integrity. This advantage positions the system as an ideal tool for preventive conservation and digital curation, particularly in regions where access to artefacts is limited due to political, climatic, or institutional constraints. Museums could employ the Model to generate reconstructions of damaged artefacts, enabling visitors to visualize their original appearance without endangering the originals.

Furthermore, the diffusion-based pipeline integrates seamlessly with 3D rendering environments, expanding its use in virtual exhibitions and educational simulations. When embedded within augmented or virtual reality (AR/VR) frameworks, the reconstructed colors can be projected onto 3D models, allowing visitors to experience artefacts in immersive chromatic fidelity. Similar integrations have been successfully implemented in the reconstruction of Roman wall paintings (Ataba, 2025) and Byzantine mosaics (Bošković and Radošević, 2019), suggesting that Harappan reconstructions could become Part of an interactive, global digital heritage narrative. In addition, the Model's modular structure allows interoperability with spectral imaging data when available. By incorporating reflectance spectra or Raman pigment profiles, the diffusion Model could serve as a hybrid reconstruction engine, balancing physical pigment evidence with generative prediction. This dual-mode functionality addresses one of the most persistent challenges in heritage colorimetry spectral-color conversion by creating a computational space where both empirical and inferential chromatic information coexist.

5.3. Heritage and ethical implications

While technological innovation enhances interpretive capacity, it simultaneously raises crucial ethical and epistemological questions. The visualization of reconstructed colors carries communicative authority; what is rendered digitally often becomes accepted as "authentic" by public audiences. To mitigate this risk, museums and researchers must adopt transparency protocols in digital color restoration.

AI-generated reconstructions should explicitly indicate their inferential nature through metadata tags, overlay visualizations, or probabilistic color maps, distinguishing between confirmed pigment zones and algorithmically reconstructed areas (ICOM, 2017). This approach aligns with the "transparency by design" principle advocated in recent digital heritage ethics frameworks (Drucker et al., 2022). Another dimension concerns cultural ownership of representation. Harappan artefacts constitute a shared South Asian heritage, yet most computational models are trained on datasets from outside the subcontinent. It is therefore critical to ensure that the reconstructed colors do not inadvertently impose cross-cultural aesthetic norms. Future research should prioritize the creation of regional pigment databases, incorporating archaeometric analyses from sites such as Mohenjo-Daro, Harappa, and

Dholavira to enhance cultural specificity.

At the same time, AI-based color restoration presents new opportunities for inclusive storytelling. The reconstructed images can be used to engage audiences, particularly younger generations, through interactive digital displays that visualize the vibrant chromatic heritage of ancient South Asia. In this sense, generative AI can act as a mediator between scientific evidence and cultural imagination, revitalizing lost visual narratives without compromising scholarly rigor. Ethically, the approach advocates a probabilistic epistemology: each color prediction represents one plausible historical scenario rather than a definitive claim. This interpretive humility not only safeguards academic integrity but also promotes dialogue between conservators, historians, and technologists about the boundaries of digital restoration.

5.4. Contribution to the special issue themes

The work presented engages directly with the special issue's focal themes by advancing the state of digital color in cultural heritage through methodological innovation, technological deployment, and interpretive reflection. Firstly, the generative Model contributes to the domain of digital and virtual rendering by offering a reproducible computational pipeline for reintegrating lost chromatic information. In doing so, it extends the remit of digital rendering from mere replication of extant surfaces towards the predictive restoration of material color, thereby aligning with contemporary trajectories in heritage 3D and virtual modelling. Secondly, by implementing a color-calibration workflow based on CIE LAB conversion and ΔE metrics, the research addresses the challenge of color acquisition and spectral-to-perceptual conversion. Although the input data were primarily high-resolution RGB images, the methodological design anticipates integration with multispectral or hyperspectral datasets, enabling future coupling of empirical pigment data with generative inference. Thirdly, the probabilistic modelling of pigment inference implicitly contributes to the debate on metameric mismatch and conservation challenges. By framing reconstructed colors as context-driven hypotheses rather than definitive retouches, the workflow mitigates the risk of misleading visual presentation under variable illuminants and display devices. Finally, the work's compatibility with immersive rendering environments (e.g., AR/VR) and its emphasis on metadata traceability position it within the broader category of advanced technologies for heritage valorization, facilitating not only scholarly analysis but also public engagement, museum interpretation, and cross-disciplinary dissemination. In sum, the research offers a synthesis of digital color science, generative machine learning, and heritage interpretation, thus fulfilling the intertwined technical and cultural ambitions of this special issue.

5.5. Limitations and future research directions

While the proposed framework demonstrates strong potential for archaeologically informed colour reconstruction, several limitations must be acknowledged. In cases where pigment traces are entirely absent, chromatic inference relies primarily on typological and contextual analogy, which increases uncertainty. Surface staining, patination, or conservation residues may occasionally be misinterpreted as pigment cues, and reconstructed pigment boundaries can appear slightly softened in heavily eroded areas. For this reason, all reconstructed outputs were reviewed by domain experts, and implausible results were excluded from aggregate quantitative evaluation. Recognizing these constraints is essential for situating generative colour reconstruction as an interpretive aid rather than a definitive restoration and provides a foundation for targeted methodological refinement in future research.

Although this study establishes a foundation for generative color reconstruction in South Asian archaeology, several avenues for further investigation remain necessary to advance the field both technically and ethically. A priority lies in developing region-specific pigment reference databases derived from archaeometric investigations of Harappan and adjacent subcontinental contexts. Embedding empirically measured reflectance, Raman, and X-ray multispectral signatures into the generative workflow would enhance chromatic accuracy and reduce reliance on cross-cultural analogues. A second direction involves augmenting the current architecture with explainable AI (XAI) mechanisms, such as saliency maps or attention-heat visualisations, to reveal how the Model derives chromatic inferences from typological, textural, and contextual cues, thereby increasing transparency and interpretive confidence. A third area for expansion concerns the integration of visual, textual, and spatial metadata into a multimodal learning framework: here, archaeological find-spot, stratigraphic layer, typological group, and pigment annotation could serve as combined inputs, enabling the Model to predict not only likely hues but also their material-social significance.

Fourthly, embedding the reconstructed color outputs into immersive museum environments via AR/VR or digital twin applications would test their efficacy in public education and stakeholder engagement, and reveal perceptual responses to color reenactment in heritage settings. Finally, the establishment of an ethical framework for generative color restoration is imperative: future work should define standards for representing uncertainty, provenance of algorithmic output, and distinctions between measured pigment surfaces and predicted reconstructions. Collaboration among archaeologists, conservation scientists, and AI specialists is necessary to ensure that chromatic reconstructions remain culturally sensitive and scholarly robust. Collectively, these research directions can transform generative color reconstruction from an experimental prototype into an integral tool within digital heritage workflows, both preserving and re-imagining the visual vitality of ancient material culture.

6. Conclusion

The study presented here introduces an innovative generative framework for reconstructing lost pigmentations in Harappan artefacts, bridging computational color science and archaeological interpretation. By combining diffusion-based image-to-image translation with calibrated colorimetry, this research demonstrates how artificial intelligence can move beyond conventional digital imaging toward a predictive understanding of chromatic loss. The results confirm that generative AI, when trained on contextually relevant datasets and evaluated through both quantitative metrics (ΔE and SSIM) and expert appraisal, can generate reconstructions that are not only visually persuasive but also archaeologically plausible.

This experiment redefines the relationship between technology and heritage: color, long regarded as an aesthetic or symbolic attribute, becomes an epistemic clue that can be algorithmically recovered. The reconstructed hues, particularly in terracotta and pottery samples, illustrate the potential of machine learning to infer material knowledge from incomplete visual evidence. The integration of a colorimetric validation step anchored in ΔE2000 thresholds ensures that AI-generated outcomes can be systematically assessed and compared against measurable pigment standards, thereby reinforcing the scientific credibility of digital restoration.

Equally significant is the cultural and ethical dimension of this work. The reconstructed images are not presented as certainties but as probabilistic visual hypotheses, guided by material, typological, and contextual logic. Such an approach reflects a broader methodological shift in digital heritage from static replication to interpretive simulation, acknowledging both the creative potential and epistemic responsibility of artificial intelligence in heritage visualization. The framework proposed here advocates transparency: every reconstruction includes embedded metadata describing the Model's assumptions, the dataset's origin, and confidence levels, thereby supporting reproducibility and interpretive accountability. Beyond its immediate archaeological context, this research establishes a scalable template for digital conservation in regions where pigment degradation or data scarcity impede traditional analytical methods (UNESCO, 2015). The Model's architecture and workflow can be adapted for other cultural materials such as murals, ceramics, or textile fragments across diverse chronological and geographical settings. Furthermore, integrating reconstructed color imagery into museum and virtual environments offers new pathways for audience engagement, enabling visitors to perceive artefacts not as faded relics but as re-envisioned testimonies of ancient sensory worlds.

The study also foregrounds the methodological synergies between artificial intelligence, color science, and cultural heritage fields that have often evolved in isolation. By situating generative modeling within the heritage discourse, the research contributes to an emergent paradigm of computational archaeology that values both precision and imagination. This convergence has implications for future digital heritage policies, especially in establishing frameworks for labeling, sharing, and archiving AI-generated cultural reconstructions. In conclusion, the proposed Generative AI Color Reconstruction (GAICR) framework represents a step toward a more dynamic and interdisciplinary vision of heritage science, where algorithms assist rather than replace human interpretation. The reconstruction of lost pigments through generative AI serves not only as a technological experiment but as an act of cultural restitution: restoring chromatic vitality to the Harappan past while simultaneously expanding the epistemological horizon of digital archaeology. Through such hybrid methodologies, the boundaries between empirical analysis and creative inference can be productively reimagined, transforming digital color from a visual attribute into a conduit for historical understanding.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work, the author(s) used ChatGPT (OpenAI) in order to improve the English grammar, clarity, and readability of the manuscript. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

CRediT authorship contribution statement

Muhammad Nouman Akhtar: Conceptualization, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Saverio Iacono:** Methodology, Supervision, Validation, Writing – review & editing. **Gianni Viardo Vercelli:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing. **Hina Zulifiqar:** Data curation, Investigation, Software, Validation, Writing – review & editing. **Muhammad Suffian:** Data curation, Formal analysis, Investigation, Resources, Validation, Writing – review & editing.

Funding

No specific funding was received for this research from any public, commercial, or not-for-profit funding agencies.

Declaration of competing interest

Authors declare they don't have any conflict of interest.

References

Amadeus, M., Castañeda, W.A.C., Zanella, A.F., Mahlow, F.R.P., 2024. From pampas to pixels: fine-tuning diffusion models for Gaúcho Heritage. arXiv Preprint arXiv: 2401.05520. https://doi.org/10.48550/arXiv.2401.05520.

Amiri, M.M., Messinger, D., Hanneken, T., 2024. Colorimetric characterization of multispectral imaging systems for visualization of historical artifacts. J. Cult. Herit. 68, 136–148. https://doi.org/10.1016/j.culher.2024.05.014.

ASTM, 2018. Standard Practice for Computing the Colors of Objects by Using the CIE System (ASTM E308–18). ASTM International. https://doi.org/10.1520/E0308-18.

Ataba, O., 2025. Digital futures of the past: AI-driven cultural heritage preservation and restoration for Africa. Zenodo. https://doi.org/10.5281/zenodo.16790241.

Barbero-Álvarez, M., Menéndez, J.M., Rodrigo, J., 2020. An adaptive colour calibration for crowdsourced images in heritage preservation science. IEEE Access 8. https://doi.org/10.1109/ACCESS.2020.3017390, 1–1.

Bošković, M.L., Radošević, N., 2019. Conservation of roman wall painting fragments from Issa: research and reconstruction of the original scheme. Retrieved from. https://api.semanticscholar.org/CorpusID:204893294.

Bolognesi, M., 2020. Virtual restoration and color reconstruction in archaeology. Digital Applications in Archaeology and Cultural Heritage 17, e00149. https://doi.org/10.1016/j.daach.2020.e00149.

Bureau of the Federal Agencies Digitization Guidelines Initiative, 2023. Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files. FADGI. https://www.digitizationguidelines.gov.

Cazzaro, I., 2025. Digital 3D Reconstruction as a Research Environment in Art and Architecture History: Uncertainty Classification and Visualisation, vol. 7. arthistoricum.net; Heidelberg University, pp. 1–392.

Chun, H., Lee, S., Sohn, J., 2025. Degradation mechanisms of traditional pigments used in historic artworks within national heritage temple halls. Kor. J. Chem. Eng. 42. https://doi.org/10.1007/s11814-025-00518-w.

ColorLab (NTNU). (n.d.). Color and Visual Computing Laboratory – NTNU. Norwegian University of Science and Technology. https://www.ntnu.edu/colorlab.

CompVis, 2022. Stable diffusion v1-5 model card. Stability AI. https://huggingface.co/runwayml/stable-diffusion-v1-5.

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26 (3), 297–302. https://doi.org/10.2307/1932409.

Drucker, J., Besser, H., McPherson, T., 2022. Ethical frameworks for digital heritage: transparency by design. Digital Scholarship in the Humanities 37 (4), 921–939. https://doi.org/10.1093/llc/fqab012.

Fadli, A., et al., 2022. AI-assisted reconstruction in cultural heritage: a review. Digital Applications in Archaeology and Cultural Heritage 27, e00231.

Fang, L., Han, B., Bi, M., Wang, L., Wang, D., 2025. DCAF-GAN: enhancing historical landscape restoration with dual-branch feature extraction and attention fusion. PLoS One 20, e0334532. https://doi.org/10.1371/journal.pone.0334532.

Fermo, P., Lombardi, C., D'Amato, A., Guglielmi, V., Giudici, B., Tomaino, A., Pozzi, M., Comite, V., Bergomi, A., Guardiano, L., Piacentini, P., 2024. Disclosing colors and pigments on archaeological objects from the Aga Khan Necropolis (West Aswan, Egypt) through on-site analytical methods: preliminary results. Heritage 7 (9), 4980–4996. https://doi.org/10.3390/heritage7090235.

Gage, J., 2009. Color and Meaning: Art, Science, and Symbolism. University of California Press.

Gallo, G., 2022. AI in heritage conservation: visual inference and virtual preservation. Heritage Science 10 (1), 87. https://doi.org/10.1186/s40494-022-00687-4.

Giesriegl, A., Pacher, U., Nagy, T., Pichler, B., Kautek, W., 2023. Laser diagnostics and processing of historical and artificial copper patina. J. Cult. Herit. 62, 45–53. https://doi.org/10.1016/j.culher.2023.05.004.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., et al., 2014. Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems, 27, pp. 2672–2680, Curran Associates, Inc. https://doi.org/10.48550/arXiv.1406.2661.

Harth, A., 2024. The study of pigments in cultural heritage: a review using machine learning. Heritage 7 (7), 3664–3695. https://doi.org/10.3390/heritage7070174.

Ho, J., Jain, A., Abbeel, P., 2020. Denoising diffusion probabilistic models. Adv. Neural Inf. Process. Syst. 33, 6840–6851.

Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, L., Chen, W., 2021. LoRA: Low-Rank adaptation of large Language models. arXiv preprint arXiv:2106.09685.

ICC (International Color Consortium), 2022. Specification ICC.2:2022 – ICC profile format. https://www.color.org/specification/ICC2v2022.pdf.

International Council of Museums (ICOM), 2017. ICOM Code of ethics for museums. https://icom.museum/en/resources/standards-guidelines/code-of-ethics.

Jaccard, P., 1912. The distribution of the flora in the alpine zone. New Phytol. 11 (2), 37–50.

Jaramillo, P., Sipiran, I., 2024. Cultural heritage 3D reconstruction with diffusion networks. arXiv Preprint arXiv:2410.10927. https://doi.org/10.48550/arXiv.2410.10927.

Krishnan, K., 1992. An analysis of decorative pigments and slips on harappan pottery from Gujarat, 17 (1), 99–108.

Liu, X., Chen, Y., Wang, Z., 2023. Diffusion models for image restoration and colorization. IEEE Trans. Image Process. 32, 1443–1456. https://doi.org/10.1109/TIP.2023.3245678.

Loshchilov, I., Hutter, F., 2019. Decoupled weight decay regularization. In: International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.1711.05101.

López-Martínez, T., Calero-Castillo, A., Carrasco-Huertas, A., Medina, V., 2024. Virtual Reconstruction Applied to Mural Painting Reintegration Projects.

Machidon, O.M., Ivanovici, M., 2018. Digital color restoration for the preservation of reversal film heritage. J. Cult. Herit. 33, 181–190.

Maietti, F., Zattini, A., 2020. Colors documentation and processing in graphic production for architectural heritage representation. In: Colour and Colorimetry: Multidisciplinary Contributions, vol. 16. Gruppo del Colore–Associazione Italiana Colore, pp. 27–34.

Merkel, D., 2014. Docker: Lightweight Linux containers for consistent development and deployment. Linux J. 2014 (239), 2.

Micheluz, A., Angelin, E.M., Lopes, J.A., Melo, M.J., Pamplona, M., 2021. Discoloration of historical plastic objects: new insight into the degradation of β-naphthol pigment lakes. Polymers 13 (14), 2278. https://doi.org/10.3390/polym13142278.

Molada-Tebar, A., Marqués-Mateu, Á., Lerma, J.L., 2019. Correct use of color for cultural heritage documentation. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences IV-2/W6, 107–113. https://doi.org/10.5194/isprs-annals-IV-2-W6-107-2019.
Muñoz-Pandiella, I., Andujar, C., Cayuela, B., Pueyo, X., Bosch, C., 2023. Automated digital color restitution of mural paintings using minimal art historian input. Comput. Graph. 114, 316–325. https://doi.org/10.1016/j.cag.2023.06.021.
OpenColorIO, 2023. OpenColorIO v2.3 documentation. https://opencolorio.org.
Paszke, A., et al., 2019. PyTorch: an imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. 32, 8024–8035.
Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., Efros, A.A., 2016. Context encoders: feature learning by inpainting. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2536–2544.
Picollo, M., Cucci, C., Casini, A., Stefani, L., 2020. Hyper-spectral imaging technique in the cultural heritage field: new possible scenarios. Sensors 20 (10), 2843. https://doi.org/10.3390/s20102843.
Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B., 2022. High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 10674–10685. https://doi.org/10.1109/CVPR52688.2022.01042.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention 9351, 234–241.
Schuhmann, C., et al., 2022. LAION-5B: an open large-scale dataset for training next generation image-text models. NeurIPS Datasets and Benchmarks. https://doi.org/10.48550/arXiv.2210.08402.
Sharma, G., Wu, W., Dalal, E.N., 2005. The CIEDE2000 color-difference formula: implementation notes and observations. Color Res. Appl. 30 (1), 21–30. https://doi.org/10.1002/col.20070.
Shoval, S., Gilboa, A., 2016. pXRF analysis of pigments in decorations on ceramics in the East Mediterranean: a test-case on Cypro-Geometric and Cypro-Archaic bichrome ceramics at Tel Dor, Israel. J. Archaeolog. Sci.: Rep. 7, 472–479. https://doi.org/10.1016/j.jasrep.2015.08.011.
Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content. Biol. Skr. 5, 1–34.
Tamburini, D., 2022. Pigments and Dyes in Archaeological and Historical objects—Scientific Analyses and Conservation Challenges. MDPI. https://doi.org/10.3390/books978-3-0365-3135-9.
UNESCO, 2015. Recommendation Concerning the Preservation Of, and Access To, Documentary Heritage Including in Digital Form. UNESCO, Paris.
Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612. https://doi.org/10.1109/TIP.2003.819861.