

Generative Artificial Intelligence for Historical Image Reconstruction Enhancing Visual Interpretation in History Education Programs

Le Thanh Hoang Tung

HUS High School for the Gifted, Hanoi, Vietnam

doi: <https://doi.org/10.37745/ejcsit.2013/vol14n7113>

Published October 11, 2026

Citation: Tung L.T.H. (2026) Generative Artificial Intelligence for Historical Image Reconstruction Enhancing Visual Interpretation in History Education Programs, *European Journal of Computer Science and Information Technology*, 14(7),1-13

Abstract: *Historical Educational engagement. There is a void in learner-focused historical reconstruction since current restoration methods seldom suit educational visualization objectives. To improve visual interpretation in history education programs, this research attempts to create a generative artificial intelligence model that reconstructs faded or damaged historical material. The suggested technique employs the Historical Heritage Painting Collection dataset, having 2,192 images. Data preprocessing incorporates noise reduction using a non-local means (NLM) denoising technique, alongside contrast normalization and scaling for uniformity. To maintain original object boundaries and direct reconstruction integrity, edge-based mapping is used for structural feature extraction. The suggested Dynamic Vortex Optimizer-driven Stable Diffusion with ControlNet (DVO-SD-CN) model is then fed the processed pictures and generated structural maps. Here, ControlNet conditions Stable Diffusion on extracted structural features to guaranty that the generated content stays true to the original historical context rather than creating irrelevant details, while DVO, a vortex-inspired metaheuristic technique, adaptively adjusts the diffusion model's hyperparameters and denoising trajectory for stable convergence. Accurate, structure-aware historical image reconstruction is made possible by this integrated architecture, making it appropriate for interactive learning modules, digital storytelling, and classroom presentation. The proposed DVO-SD-CN achieves PSNR values of 33.84, 32.47, and 29.86 for low, moderate, and high deterioration, respectively, according to experimental evaluation. All things considered, the suggested approach shows great promise for improving visual interpretation in history education programs using dependable, technologically advanced picture reconstruction techniques.*

Keywords: generative artificial intelligence, historical image reconstruction, stable diffusion, controlnet, visual interpretation, history education.

INTRODUCTION

Historical education uses a wide range of textual and visual resources in form of texts, historical documents, image, paintings, maps, and various images to describe events, individuals, places, cultures, and sociological shifts within the different historical periods. The use of visual materials is especially helpful as it allows making easier for students to get familiar with the historical environment and draw better associations with the events and their context (Jo et al., 2024). Nevertheless, numerous historical images have been corrupted, lost, destroyed because of natural degradation, poor protection, or destruction of the original sources. Consequently, the lack of authentic visual materials can create

difficulties for students to obtain a complete vision of historical places, artefacts, architectures, and cultural environment. Digital archives and computational technologies allow overcoming the described problems and using historical materials efficiently (Liu et al., 2024). An additional solution in terms of historical education can be provided with the use of Generative Artificial Intelligence (AI). It can assist in restoration and visualization of incomplete or degraded historical imagery. Reconstruction of historical visuals with the use of AI can complement archival and textual material with the visual representations of the historical buildings, landscapes, artefacts, cultural environments, and other historically relevant subjects. The suggested methodology uses AI, Machine Learning (ML) and Deep Learning (DL) approaches for the purposes of historical imagery reconstruction and visualization. First, the images and data related to them are extracted from digital archives, repositories, and educational materials. Then ML algorithms can be used for classification, recognition of the patterns, evaluation of the similarity, and categorization of historical imagery (Dong et al., 2025). Next, DL methods can be applied for the extraction of complex visual features of historical images including shapes, textures, structure, and contextual characteristics. The reconstructed images can be produced on the basis of the obtained visual features and historical information about the image subject. The textual descriptions of the historical subjects can also be used as the contextual information for the reconstruction. The resulting images can be used further in educational resources to facilitate the understanding of historical events, places, objects, buildings, and cultural environment. The proposed approach is especially useful in cases, when historical imagery is degraded, corrupted, and incomplete or totally absent. The textual information cannot be enough to create the vision of historical settings, artefacts, architectural structures, and events (Fareed et al., 2024). Conventional image restoration and reconstruction require considerable human involvement, expertise, and time. The existing image enhancement technologies allow improving the quality of degraded images. Nevertheless, such technologies do not use historical information effectively when the image is restored. Thus, an intelligent restoration approach is needed to use both historical visual information and relevant textual context.

1.1. Research Objective and Contribution

This research designs a Generative AI for reconstructing historical images for better visual interpretation in history education programs. The contributions of this research are listed below:

- The proposed DVO-SD-CN uses AI and generative modeling to reconstruct damaged, faded, and incomplete historical images.
- Combines image preprocessing, structural feature extraction, optimization through the DVO algorithm, and stable diffusion through ControlNet to preserve visual structures and improve the contextual quality of historical images.
- Offers a practical method for the use of reconstructed historical images in classroom visualization, digital storytelling, and learning to make visual interpretation of historical events, objects, and environments possible.

Research Organization

Remaining research is grouped into four areas: Section 2 covers existing methodologies for historical image restoration, generative image reconstruction, and AI-assisted visual learning while outlining the research need. Section 3 provides the proposed technique, including dataset collection, preprocessing, structural feature extraction, and the DVO-SD-CN. The experimental findings, comparison analysis, picture quality evaluation, and instructional assessment of the restored historical images are presented in Section 4. Research is concluded in Section 5, which summarizes the main conclusions and discusses possible future development paths.

RELATED WORK

Previously done research has increasingly examined the use of Generative AI to reconstruct, restore, and visualize historical and cultural heritage images. The main focus of the existing literature is on the generation of relevant images, reconstruction of missing parts, restoration of heritage architecture, reproduction of damaged artifacts, as well as recovery of missing colors or pigments. Various methods have been considered, such as prompt-based text-to-image generation, image inpainting using AI, GAN-based restoration, and the use of historical data as guidance for generative models. While some works rely on visual assessments or perceptual similarity estimates provided by participants, others use various metrics such as accuracy, ΔE_{00} , PSNR, SSIM, LPIPS, and FID. This evolution shows the development from subjective visual evaluations to a more complete quantitative analysis.

Table 1: Related Works

Ref.	Aim	Method	Result	Conclusion	Limitation
Díaz-Sánchez and Chapinal-Heras (2024)	To investigate the use of Open Access Generative AI for visualizing Classical Antiquity.	Used MidJourney with descriptions and prompts to generate visual representations of the ancient past.	Generated visualizations that could be used to interpret Classical Antiquity.	Open-access Generative AI can support the visualization and interpretation of historical subjects.	Quantitative restoration parameters were not emphasized.
Richtr (2025)	To reconstruct missing parts of images from the Czech Republic.	Applied AI-based image inpainting to reconstruct missing image regions.	Produced reconstructed images that were primarily assessed visually.	AI-based inpainting can support the reconstruction of missing historical image content.	Evaluation was mainly visual rather than quantitatively measuring reconstruction fidelity.
Arzoman d et al. (2024)	To restore architectural heritage using historical information and visual guidance.	Used text-to-image Generative AI, historical data, visual guidance, and prompt engineering.	Generated reconstructed representations of architectural heritage.	Combining historical information with visual guidance can improve AI-based architectural reconstruction.	Image fidelity was measured quantitatively only to a limited extent.
Sukkar et al. (2024)	To evaluate AI-generated representations of Islamic architectural features,	Generated architectural representations using AI and evaluated their similarity through	Similarity scores ranged from 50.0% to 73.9% across different minaret types.	AI can generate visually recognizable representations of historical	Evaluation depended mainly on participant-based similarity assessment rather than comprehensive objective metrics.

	particularly minarets.	participant assessments.		Islamic architecture.	
Altaweel et al. (2024)	To restore damaged historical coins and evaluate their authenticity.	Applied Generative Adversarial Networks (GANs) for coin restoration.	Achieved an average 46% accuracy in authenticity discrimination.	GAN-based approaches can support the restoration of damaged historical artefacts.	Authenticity discrimination accuracy remained limited, indicating challenges in reliably validating reconstructed content.
Akhtar et al. (2026)	To restore lost pigments in Harappan historical materials.	Applied Generative AI for pigment restoration and evaluated colour and structural reconstruction.	Achieved $\Delta E_{00} = 3.67$ for colour restoration and >0.95 for structural similarity.	Generative AI can support quantitative restoration of lost historical colours and structural information.	The study focuses primarily on pigment restoration, leaving broader visual reconstruction and multimodal historical context less addressed.
DCAF-GAN	To evaluate generative image restoration using comprehensive image-quality measures.	Used a Dual-branch Convolutional Attention Fusion GAN (DCAF-GAN) on StreetView and Places2 datasets.	StreetView: PSNR 29.12 dB, SSIM 0.867, LPIPS 0.137, FID 14.4. Places2: PSNR 28.6 dB, SSIM 0.854, LPIPS 0.143, FID 15.3.	Combining multiple quantitative metrics provides a broader assessment of reconstruction quality.	The datasets are general image datasets rather than specifically historical or cultural heritage collections.

Research gap

Generative AI has shown the ability to help in some areas of visualization of historical and cultural heritage. Díaz-Sánchez and Chapinal-Heras (2024) researched the majority of current generative AI research focuses on certain legacy domains, including historical landscapes, coinage, architecture, and archeological pigments. There hasn't been enough research done on a general framework for rebuilding various deteriorated or fragmentary historical photographs. Richtr (2025) research the application of AI for texture completion in damaged heritage images. Conventional text-to-image creation can yield visually realistic visuals but may not preserve the original structural properties of historical information. Arzomand et al. (2024) research the reconstruction of historical architecture using historical data, visual cues and prompt engineering. Adaptive optimization for generative-model parameters is only partially taken into account by current heritage reconstruction techniques. Coordinated Investigation of structural Reconstruction Coordinated. Sukkar et al. (2024) analyzed AI-generated representations of Islamic architectural heritage focused on the restoration of damaged cultural artefacts, namely historical coins, using generative AI techniques. Akhtar et al. (2026) Research the digital restoration of lost colors in archaeological heritage while generative AI has been investigated for heritage education and restoration, there hasn't been much focus on combining educational visualization with reconstructed historical imagery. It is consequently necessary to have a controlled reconstruction framework that supports educational content that is both visually understandable and historically knowledgeable. Fang et al.

(2025) concentrated on the restoration of historical landscapes. Moreover, the issue of explicit structural conditioning and adaptive parameter optimization, the proposed DVO-SD-CN approach combines Stable Diffusion with ControlNet and DVO. ControlNet uses edge-based structural conditioning to direct the process of reconstruction according to the structural information available in historical images. DVO adaptively optimizes the set of key parameters of Stable Diffusion and ControlNet. Thanks to the combination of historical visual information, structural guidance and adaptive parameter optimization, DVO-SD-CN is aimed at providing more controlled and context-aware historical image reconstruction for educational visualization.

METHODOLOGY

The proposed DVO-SD-CN involves data preprocessing, structural feature extraction, generative reconstruction, and adaptive optimization. While Sobel edge detection retrieves important edges, contours, and structural characteristics from historical image, denoising utilizing Non-Local Means, contrast normalization, and image scaling improves their quality and consistency. In order to direct Stable Diffusion and preserve the spatial characteristics of the input heritage images, ControlNet is subsequently fed these structural maps. Using a fitness function of SSIM, LPIPS, and normalized PSNR, the DVO adaptively optimizes a number of reconstruction parameters, including denoising intensity, guidance scale, number of sampling steps, and strength of ControlNet guidance. The entire process of the suggested DVO-SD-CN is described in Fig. 1.

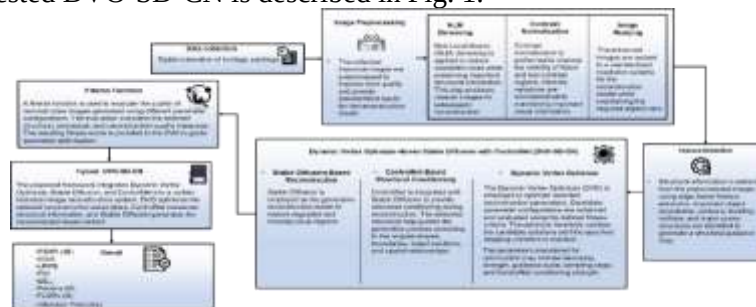


Fig.1. Proposed DVO-SD-CN framework for digital restoration of historical paintings.

Data collection

Historical Heritage Painting Collection collection consists of 2,192 images of paintings with historical relevance from prominent institutions. It is a collection of paintings from various eras with varying degrees of deterioration, including fading, discoloration, cracking, and other types of damage. The images were gathered from the Metropolitan Museum of Art, Harvard University's Art Museum, Princeton University's Art Museum, and the Smithsonian's Freer Gallery of Art. A proportionate 80:20 split was used to separate the training and testing datasets from the data. So, 1,754 images (80%) are in the training set, while 438 images (20%) are in the testing data. Therefore, 2, this dataset comprises attributes linked to images, such as Image ID, Museum Source, Image Width, Image Height, Brightness, Contrast, Sharpness, Edge Density, Dynamic Range, Saturation, Image Quality Score, Degradation Score, and Target. There are three types of deterioration for the target attribute: mild degradation (0), moderate degradation (1), and high degradation (2). The dataset used here is collected from **Kaggle**: <https://www.kaggle.com/datasets/colabsss/historical-heritage-painting-collection>

Data preprocessing

The first step is the preparation of the preparation of the preparation of the preparation of the process of the structural visual pictures of the structural reconstruction of the input database. The preprocessing step contains NLM denoising, contrast normalization, and image scaling.

Non-Local Means (NLM) Denoising: The structural and visual elements of the paintings that are part of cultural heritage, such as architecture, object borders, texture, and artistic details, are preserved while the noise is eliminated using the NLM denoising algorithm in Equation (1).

$$I(p) = \frac{1}{\sum_{q \in \Omega} \exp \left(- \frac{\|Pp - Pq\|^2}{h^2} \right)} I(q) \quad (1)$$

Here, Pp and Pq are image patches at positions $(I(p))$. The strength of the filter is controlled by parameter h^2 , while $\sum_{q \in \Omega} \exp Z(p)$ is the normalization term.

Contrast Normalization: In historical cultural heritage image, the contrast normalization technique is used to highlight the low-contrast and faded parts. The visual information is preserved while the intensity measurements are adjusted into a standard rang in Equation (2).

$$Icn(x, y) = \frac{I(x, y) - Imin}{Imax - Imin} \quad (2)$$

Where $I(x, y)$ is the intensity of the input image and $Imin$ and $Imax$ are the minimum and maximum $Icn(\quad)(x, y)$ image intensities, respectively.

Image Resizing: After normalization of contrast in cultural heritage images, they are scaled to a common size suitable for the generative reconstruction algorithm. Bilinear interpolation is applied to derive the I_r intensity at the desired spatial position in Equation (3).

$$I_r(x, y) = \sum_{i=0}^1 \sum_{j=0}^1 w_{ij} I(x_i, y_j) \quad (3)$$

Here w_{ij} stands for the interpolation weights, while $I(x_i, y_j)$ is the resized image. $\sum_{i=0}^1$ Considers two adjacent values of (i) : $(i = 0)$ and $(i = 1)$. $\sum_{j=0}^1$ Considers two adjacent values of (j) : $(j = 0)$ and $(j = 1)$.

Edge-based Structural Information Extraction

To direct the generative reconstruction process, structural information is taken from the preprocessed images of the cultural heritage collection. Important borders, contours, architectural features, items, and scene structures are identified via edge-based extraction. ControlNet is then used to convert the retrieved structural data into guidance maps for conditioning. The Sobel operator is used to calculate image gradients in both horizontal and vertical directions. The degree of structural edges present in the historical cultural heritage environment is indicated by the value of the gradient magnitude $G(x, y)$ in Equation (4).

$$G(x, y) = (I * K_x)^2 + (I * K_y)^2 \quad (4)$$

Where $(I * K_x)^2$ and $(I * K_y)^2$ represent the horizontal and vertical Sobel masks, respectively.

3.4 DVO-SD-CN for Structure-Guided Hybrid Reconstruction

The proposed DVO-SD-CN model is an integrated model that reconstructs deteriorated cultural heritage images using Stable Diffusion, ControlNet, and DVO. As a generative reconstruction module, Stable Diffusion can reconstruct degraded regions while producing realistic visual data. In the meantime, ControlNet functions as a structural conditioner by utilizing the edges, contours, object boundaries, and primary structures of the preprocessed cultural heritage images. Conditioning makes ensuring that the visual material produced is in line with the images' spatial organization and structural characteristics. On the other hand, the DVO optimizes the reconstruction parameters such as denoising strength, guidance scale, sampling steps, and ControlNet conditioning strength depending on the candidate configurations of the parameters evaluated by the fitness function.

Stable Diffusion-based Image Reconstruction: The generated Diffusion is the generated Diffusion is a progressive Diffusion is the generated. In the reconstruction stage, the model learns and generates visual aspects including texture, shape, color, and context to reconstruct the degraded or missing parts of the cultural heritage paintings. Stable Diffusion serves as the generator in the proposed DVO-SD-

CN model, while ControlNet's structural information impacts the reconstruction to match the spatial components of the original image in Equation (5).

$$x_t = \sqrt{\alpha_t}x_0 + \sqrt{1 - \alpha_t}\epsilon, \epsilon \sim N(0, I) \quad (5)$$

Where x_0 is the original image representation, x_t is the noisy representation at step t , $\sqrt{\alpha_t}x_0 + \sqrt{(1 - \alpha_t)}\epsilon$ is the noise scheduling parameter, and $\epsilon \sim N(0, I)$ is the Gaussian noise.

Structural Conditioning through ControlNet: A structural guidance map is created from a preprocessed cultural heritage image utilizing edge-based feature extraction, capturing significant spatial features such as object borders and architectural features. This is the spatial layout that was produced and functions as spatial content. ControlNet leverages the guidance map to retain structural integrity, thereby maintaining crucial boundaries and spatial relationships in the historical image reconstruction, whereas Stable Diffusion concentrates on providing visual content and filling in missing sections using Equation (6).

$$\epsilon\theta = \epsilon\theta(xt, t, c, S) \quad (6)$$

Where $\epsilon\theta$ xt is the noisy encoding, t is the diffusion step, c refers to the condition for reconstruction, and S is the structural guidance map.

Dynamic Vortex Optimizer: When it comes to adjusting parameters for the reconstruction process utilizing the Stable Diffusion-ControlNet architecture, the traditional parameter tuning technique has certain drawbacks. While random search may lack systematic examination of the regions with good potential, grid search would require the evaluation of several configurations of given parameters, resulting in a large computing cost. Another difficulty in optimizing the non-linear and non-convex function of reconstruction is gradient-based optimization. DVO is used to dynamically identify optimal configurations in order to overcome these constraints. It may select the best reconstruction parameters X_i for enhanced image quality while maintaining the structural details of cultural heritage paintings using such an adaptive optimization technique. using Equation (7).

$$X_i = [d_i, g_i, s_i, c_i] \quad (7)$$

Where d_i , g_i , s_i , and c_i are denoising strength, guidance scale, number of sampling steps, and ControlNet conditioning strength, respectively.

Fitness Evaluation: The fitness of each configuration is estimated through structural and perceptual assessments of the reconstruction quality. This F fitness value determines the proper parameter configuration for the historical cultural heritage reconstruction by the DVO in Equation (8).

$$F = w_1(1 - SSIM) + w_2(1 - LPIPS) + w_3(1 - PSNR_n) \quad (8)$$

Where $1 - SSIM$ stands for structural similarity, $1 - LPIPS$ stands for perceptual similarity, $1 - PSNR_n$ stands for reconstruction quality normalization, and w_1, w_2, w_3 are the weights, respectively. The hybrid system Preprocessing, feature extraction, Stable Diffusion, ControlNet, and DVO while contrast normalization enhances the visibility of faded and low-contrast areas of an image, NLM denoising removes undesired noise from an image. The input preprocessed images are evaluated to extract structural elements such as edges, contours, borders, and spatial information. Control is supplied guidance While ControlNet maintains the structural characteristics of the original painting, the reconstructed image produced by the Stable Diffusion algorithm replaces the damaged or absent visual elements of an image. The DVO reconstructing the hybrid system is the hybrid system. Thus, the hybrid system (Algorithm 1) creates a sequence of steps, in which preprocessing increases the quality of the input, feature extraction generates spatial guidance, Stable Diffusion carries out image reconstruction, ControlNet preserves structural information, and DVO optimizes the reconstruction process. The images obtained after the reconstruction process are useful for visual analysis and the presentation of historical content in history education.

Table 2: Algorithm 1

Algorithm 1: DVO-SD-CN-Based Historical Cultural Heritage Image Reconstruction*Input: Historical cultural heritage image I Output: Reconstructed image $I^* I^{\wedge}$*

1. *Start*
2. *Read historical cultural heritage image I .*
3. *Apply NLM denoising to obtain I_{dI} .*
4. *Apply contrast normalization to obtain I_{cn} .*
5. *Resize I_{cn} to obtain preprocessed image I_{pI} .*
6. *Apply Sobel edge detection to I_{pI} .*
7. *Generate structural guidance map SS .*
8. *Initialize DVO population $X = \{X_1, X_2, \dots, X_N\}$.*
9. *Set iteration $t = 1$.*
10. *While $t \leq T_{max} \ t \leq T_{\max}$ do*
11. *For each candidate parameter X_i do*
12. *Set Stable Diffusion and ControlNet parameters using X_i .*
13. *Generate reconstructed image I_i .*
14. *Calculate fitness value F_i .*
15. *End For*
16. *Select the candidate with the best fitness value.*
17. *Update DVO candidate parameters.*
18. *Set $t = t + 1$.*
19. *End While*
20. *Select the optimized parameter configuration X^* .*
21. *Generate the final image I^* using Stable Diffusion and ControlNet.*
22. *Return the reconstructed cultural heritage image I^* .*
23. *End*

RESULT AND DISCUSSION

The DVO-SD-CN model was created using Python on an Intel Core i7 with 16 GB RAM and an NVIDIA GeForce RTX 3060. PSNR, SSIM, LPIPS, and FID were used to examine the model's performance.

Experimental Results

The model shows stable learning and proper convergence after 50 epochs as depicted in Fig.2 (a and b). Training loss reduces from 1.08 to 0.08 and the validation loss from 1.18 to 0.12, showing that the learning errors have been minimized. Accuracy on training is 63.5% to 98.8%, and validation accuracy is 59.8% to 96.7%. The improvement in performance in training and validation shows that the model learns visual attributes from the historical paintings dataset effectively. This result helps to achieve the goal of building a robust framework for reconstructing historical images.

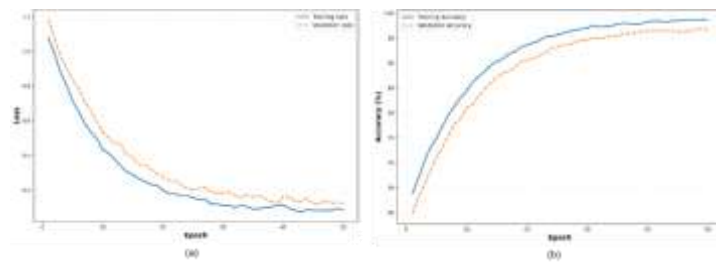


Fig.2. Analysis of (a) Loss curves for training and validation phases during 50 epochs. (b) Accuracy curves for training and validation phases during 50 epochs.

The model shows excellent classification and reconstruction performance for all the classes. For example, in Fig 3(a), the curves show the discriminative power with Area Under the Curve (AUC) values of 0.9845 for class 0, 0.9846 for class 1, and 0.9949 for class 2, as compared to the baseline model. Fig 3 (b) confirms the above results with precision-recall curves, giving an Average Precision (AP) score of 0.9747, 0.9575, and 0.9888 for classes 0, 1, and 2, respectively. Image restoration for historical images with varying degrees of degradation, ranging from low to high, while keeping structural, perceptual, and color integrity intact.

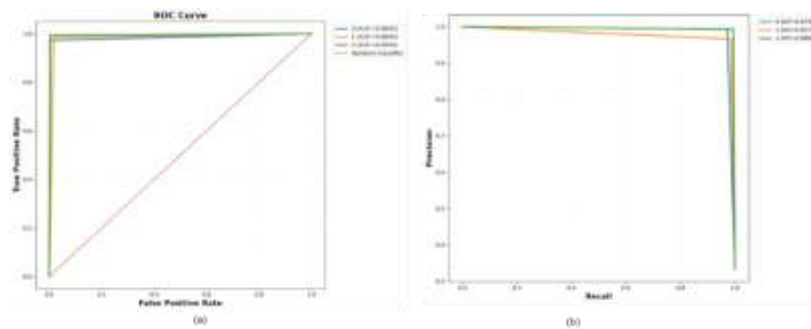


Fig.3. Performance of: (a) Curves that show evaluation for classes 0, 1, and 2. (b) Curves that show precision and recall for classes 0, 1, and 2.

The confusion matrix in Fig.4 measures the accuracy of classification for each of the three classes of degradation. For Class 0, there are 141 true positives and 4 false positives belonging to Class 1. In the case of Class 1, there are 143 true positives, 1 false negative belonging to Class 0, and 1 false negative belonging to Class 2. Finally, in Class 2, there are 147 true positives and 1 false positive belonging to Class 1. For image restoration of historical images, the purpose of this research is to determine whether the model can effectively distinguish between low, medium, and high degrees of degradation.

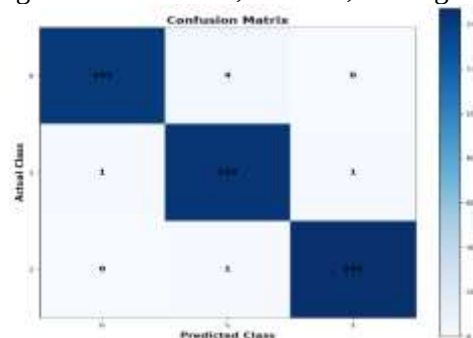


Fig.4. Confusion Matrix assessing the multi-class classification results for class labels 0, 1, and 2.

Preprocessing Pipeline Analysis

The preprocessing pipeline standardizes historical visual inputs before feeding them to the generative models. Fig. 5 (a) shows the preprocessing pipeline: denoising using NLM, normalization of contrast, and resizing to 256 x 256. In Fig.5 (b) it goes through the same 4-step preprocessing pipeline. In Fig.5 (c), it goes through cleaning and rescaling to 256 x 256. In Fig.5 (d), it goes through the same preprocessing steps. In Fig.5 (e), it goes through the standardized preprocessing process.

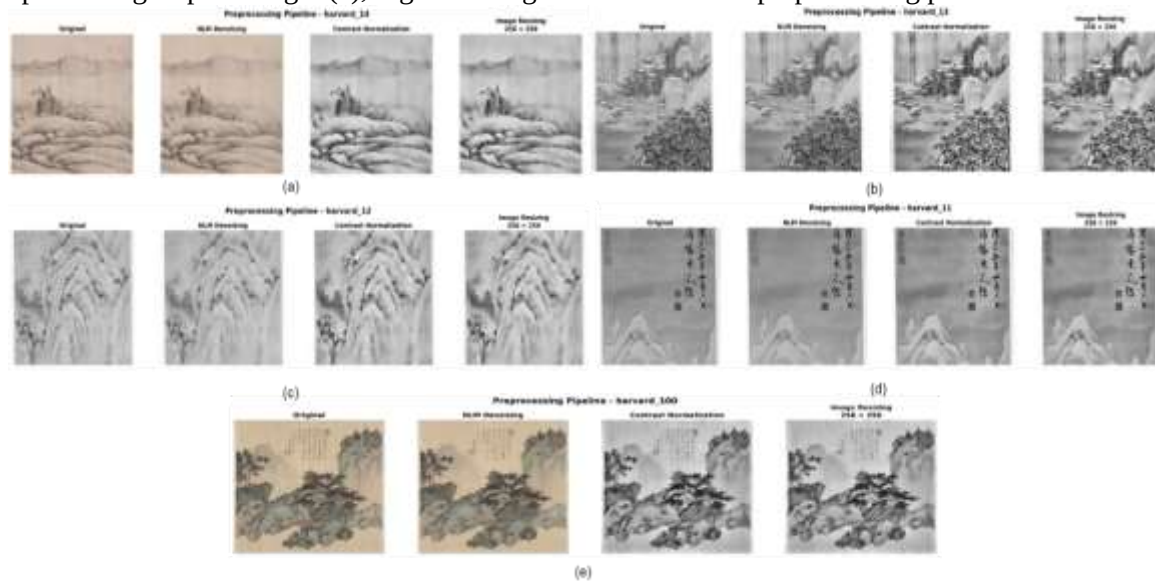


Fig.5. Preprocessing pipeline for selected cultural heritage paintings: (a) original images, (b) NLM denoising, (c) contrast normalization, (d) image resizing to 256×256 pixels, and (e) final preprocessed images.

Quantitative Feature Fusion and Extraction Analysis

In this part, extraction of quantitative features from historical painting samples is provided for Generative Artificial Intelligence for Historical Image Reconstruction and visualization in history education curriculum. In Fig.6. (a), global characteristics of the sample indicate a mean intensity of 158.17, a standard deviation of 29.63, a minimum intensity of 4, and a maximum intensity of 249, giving rise to an image quality of 0.234 and a degradation of 0.765. In Fig.6.(b), regional spatial characteristics of the same sample indicate that the mean intensity in the top-left region is 146.05 with a standard deviation of 19.86. The GLCM correlation and ASM are 0.906 and 0.072, respectively.

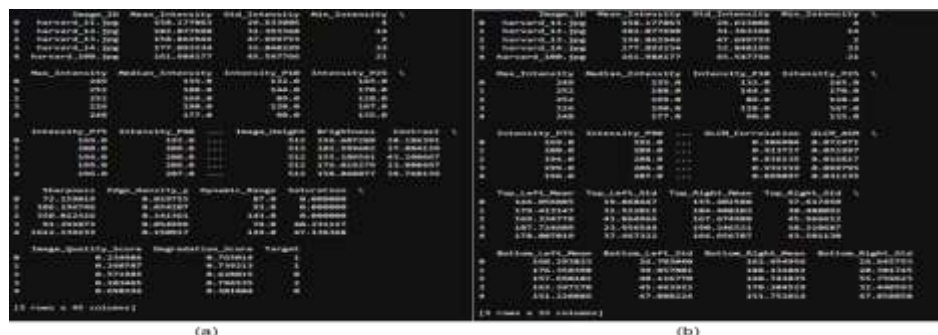


Fig.6. Analysis of (a) Global measures of intensity, quality, and degradation from historical images. (b) Measures of regional texture and spatial statistics from historical images.

Metrics Explanation

The performance of the reconstruction is measured based on PSNR, SSIM, LPIPS, FID, and ΔE_{00} . **PSNR**: a measure of how similar the reconstructed and reference images are, with larger numbers representing more similar reconstructions. Structural Similarity Index Measure (**SSIM**): a measure of structural similarity, taking into consideration luminance, contrast, and structural properties of the images, with larger numbers representing more structure-preserving reconstructions. Learned Perceptual Image Patch Similarity (**LPIPS**): a measure of perceptual similarity between reconstructed and reference images, with smaller numbers representing more perceptually similar reconstructions. Fréchet Inception Distance (**FID**): a measure of the similarity of distributions of generated images and reference images, with smaller numbers representing more similar reconstructions. Finally, CIEDE2000 Color Difference (**ΔE_{00}**): a measure of the color difference between reconstructed and reference images, with smaller numbers representing more accurate reconstructions.

Comparison evaluation

Both the existing techniques and the DVO-SD-CN technique are tested under similar conditions when dealing with historical heritage paintings' restoration in varying degrees of degradation. The existing restoration techniques include GAICR (Akhtar et al., 2026) and DCAF-GAN (Fang et al., 2025). The proposed DVO-SD-CN technique is based on a combination of Stable Diffusion, ControlNet structural conditioning, and DVO techniques. Table 1 comparison shows that the proposed technique produces more accurate results in terms of reconstruction, structures, distributions, and color preservation than the existing ones.

Table 3: Performance Comparison of GAICR, DCAF-GAN, and Proposed DVO-SD-CN

Method	Target Class	PSNR (dB)	SSIM	LPIPS	FID	ΔE_{00}	Params (M)	FLOPs (G)	Inference Time (ms)
GAICR	Low Degradation	30.12	0.9015	0.1284	26.74	6.21	32.80	118.40	42.70
	Moderate Degradation	28.36	0.8587	0.1762	35.82	8.13	32.80	118.40	42.70
	High Degradation	25.04	0.7661	0.2498	48.61	11.47	32.80	118.40	42.70
DCAF-GAN	Low Degradation	31.45	0.9182	0.1087	22.63	5.48	28.60	102.70	36.80
	Moderate Degradation	30.02	0.8864	0.1425	29.74	6.73	28.60	102.70	36.80
	High Degradation	27.42	0.8216	0.2042	40.18	9.26	28.60	102.70	36.80
DVO-SD-CN [Proposed]	Low Degradation	33.84	0.9447	0.0821	17.26	4.21	21.40	78.60	28.50
	Moderate Degradation	32.47	0.9218	0.1037	21.83	5.38	21.40	78.60	28.50
	High Degradation	29.86	0.8795	0.1516	30.47	7.12	21.40	78.60	28.50

DISCUSSION

When the degree of degradation increases, existing methods like GAICR and DCAF-GAN struggle to restore historical paintings, leading to a decline in structural preservation, visual perception, and color

coherence. By combining Stable Diffusion, structural guidance using ControlNet based on Sobel, and DVO, the DVO-SD-CN approach tackles the aforementioned problems. ControlNet assists in maintaining important boundaries, contours, and spatial relationships, while the Stable Diffusion technique restores or reconstructs the degraded visual information. In the meantime, DVO dynamically adjusts the reconstruction process to improve reconstruction consistency over a range of deterioration levels

Practical implications: Digital storytelling, online history-learning modules, digital archives of historical assets, and the presentation of deteriorated historical paintings can all benefit from the suggested approach. The student can explore visual details in reconstructed images that are difficult to see in the severely damaged original. Reconstructed outputs should be accompanied by the original photos and any available historical data, and they should be explicitly marked as AI-generated outputs.

Limitations: These reconstructed images do not represent or reflect other historical items (e.g. photos, handwritten documents, maps, sculptures, architectural blueprints, and similar items). Both generative reconstruction and structural conditions can incorporate visually plausible elements lacking historical documentation, though structural conditions do not guarantee historical accuracy. Therefore, the pedagogical implications must be confirmed through structured evaluation by both learners and educators.

CONCLUSION

Research proposed the DVO-SD-CN architecture to successfully restore degraded paintings through the integration of Stable Diffusion, ControlNet, and DVO. The architecture provided PSNR of 33.84, 32.47, and 29.86 dB and SSIM of 0.9447, 0.9218, and 0.8795 for low, moderate, and high levels of degradation, respectively. In addition, the architecture provided LPIPS of 0.0821, 0.1037, and 0.1516, FID of 17.26, 21.83, and 30.47, and ΔE_{∞} of 4.21, 5.38, and 7.12, showing improved performance in terms of perceptual, structural, distributional, and color restoration qualities. These findings show how the architecture could be used to depict history for educational and digital heritage purposes. Nonetheless, the restored paintings should be considered AI visualization of history and not necessarily a historically authentic content. For further research, a larger and wider range of data can be included, along with historical metadata, expert authenticity assessment, and classroom-based evaluation.

REFERENCES

1. Jo, H., Lee, J.K., Lee, Y.C. and Choo, S., 2024. Generative artificial intelligence and building design: early photorealistic render visualization of façades using local identity-trained models. *Journal of Computational Design and Engineering*, 11(2), pp.85–105. <https://doi.org/10.1093/jcde/qwae017>.
2. Liu, J., Ma, X., Wang, L. and Pei, L., 2024. How can generative artificial intelligence techniques facilitate intelligent research into ancient books? *ACM Journal on Computing and Cultural Heritage*, 17(4), pp.1–20. <https://doi.org/10.1145/3690391>.
3. Dong, Q., He, J., Li, N., Wang, B., Lu, H. and Yang, Y., 2025. Exploring the cognitive reconstruction mechanism of generative AI in outcome-based design education: a study on load optimization and performance impact based on dual-path teaching. *Buildings*, 15(16), p.2864. <https://doi.org/10.3390/buildings15162864>.
4. Fareed, M.W., Bou Nassif, A. and Nofal, E., 2024. Exploring the potentials of artificial intelligence image generators for educating the history of architecture. *Heritage*, 7(3), pp.1727–1753. <https://doi.org/10.3390/heritage7030081>.

5. Díaz-Sánchez, C. and Chapinal-Heras, D., 2024. Use of open access AI in teaching classical antiquity: a methodological proposal. *The Journal of Classics Teaching*, 25, pp.17–21. <https://doi.org/10.1017/S2058631023000739>.
6. Richtr, R., 2025. Photorealistic texture contextual fill-in. *Heritage*, 8(1), p.9. <https://doi.org/10.3390/heritage8010009>.
7. Arzomand, K., Rustell, M. and Kalganova, T., 2024. From ruins to reconstruction: harnessing text-to-image AI for restoring historical architectures. *Challenge Journal of Structural Mechanics*, 10(2), pp.69–85. <https://doi.org/10.20528/cjsmec.2024.02.004>.
8. Sukkar, A.W., Fareed, M.W., Yahia, M.W., Abdalla, S.B., Ibrahim, I. and Senjab, K.A.K., 2024. Analytical evaluation of Midjourney Architectural Virtual Lab: defining major current limits in AI-generated representations of Islamic architectural heritage. *Buildings*, 14(3), p.786. <https://doi.org/10.3390/buildings14030786>.
9. Altaweel, M., Khelifi, A. and Zafar, M.H., 2024. Using generative AI for reconstructing cultural artifacts: examples using Roman coins. *Journal of Computer Applications in Archaeology*, 7(1), pp.301–315. <https://doi.org/10.5334/jcaa.146>.
10. Akhtar, M.N., Iacono, S., Vercelli, G., Zulifiqar, H. and Suffian, M., 2026. Generative AI for digital color restoration in archaeology: reconstructing lost pigments of the Harappan Civilization. *Digital Applications in Archaeology and Cultural Heritage*, p.e00549. <https://doi.org/10.1016/j.daach.2026.e00549>.
11. Fang, L., Han, B., Bi, M., Wang, L. and Wang, D., 2025. DCAF-GAN: enhancing historical landscape restoration with dual-branch feature extraction and attention fusion. *PLOS ONE*, 20(10), p.e0334532. <https://doi.org/10.1371/journal.pone.0334532>.