The Polity and Society

www.tpas.lexarcheus.com
Vol.-03, No.-02 (Jul-Dec, 2025)

ISSN: 2584-0770
infotpas@gmail.com

Exploring the Intersection of Technology and History: Stable Diffusion Inpainting in Heritage Restoration

Kajol Chaudhary[1], Tanni Moitra[2]

[1] Student of MA in Modern History, Sharda School of Humanities & Social Sciences, Sharda University, India.
[2] Assistant Professor of History, Sharda School of Humanities & Social Sciences, Sharda University, India.

*** Corresponding Author:**

Kajol Chaudhary
kajolc35@gmail.com
https://doi.org/10.5281/zenodo.18278078

ARTICLE INFO

Article History

Received	01-07-2025
Revised	15-08-2025
Accepted	19-11-2025
Published	30-12-2025

ARTICLE ID

TPAS0302001

KEYWORDS

Stable Diffusion, Inpainting, Heritage Restoration, Machine Learning, and Digital Preservation.

.

OPEN ACCESS

ABSTRACT

Preserving cultural heritage is essential for sustaining historical knowledge and architectural continuity. This study examines the use of Stable Diffusion Inpainting for the digital restoration of partially damaged monuments, with focus on the gopurams of the Vittala Temple in Hampi and Salimgarh Fort in Delhi. The research applies the Runway ML v1.5 Stable Diffusion Inpainting model to reconstruct missing architectural elements preserving spatial accuracy and structural balance. The methodology involves image preprocessing through noise reduction and damaged region identification, followed by systematic model implementation and evaluation. The findings show that the model effectively restores complex forms and ornamental details with high visual coherence and minimal structural distortion. The case studies confirm practical applicability also identifying limitations related to texture alignment and historical precision. Overall, the study finds that Stable Diffusion Inpainting is a reliable digital tool that can support heritage conservation and informed restoration practices.

INTRODUCTION

Cultural heritage reflects the long journey of human civilization and records social values, artistic skills, and technological knowledge across time. Ancient monuments, temples, and forts serve as physical evidence of this legacy. According to UNESCO, more than one thousand heritage sites worldwide face moderate to severe risk due to natural decay, climate change, pollution, conflict, and rapid urbanization (UNESCO, 2021). The loss of these sites weakens cultural identity and disrupts the transmission of historical knowledge to future generations. Restoration and reconstruction of damaged heritage sites are therefore essential for maintaining historical continuity and public memory (Childers, 1979).

Restoration is not limited to structural repair. It reconnects past and present by allowing societies to understand their origins and Cultural Revolution. Traditional restoration methods depend heavily on manual expertise, archival records, and physical materials. These methods are time consuming, expensive, and often constrained by incomplete documentation (Ahmed et al., 2021). As a result, many partially destroyed monuments remain unrestored or inaccurately reconstructed.

Recent advances in machine learning offer new possibilities for addressing these limitations. Machine learning systems can analyze large volumes of visual and structural data, recognize complex patterns, and generate context aware reconstructions (Efros et al., 1999). In heritage studies, these capabilities support digital preservation, virtual reconstruction, and informed restoration planning. Studies show that AI based image restoration can significantly improve visual coherence and reduce reconstruction errors when compared to rule based digital methods (Barnes et al., 2009). Within this technological shift, Stable Diffusion Inpainting has emerged as a powerful tool for reconstructing damaged visual content. This technique can intelligently fill missing or destroyed regions of images while preserving architectural structure, spatial alignment, and stylistic consistency (Bertalmio et al., 2000). Its application to heritage restoration is particularly relevant for monuments with partial damage rather than complete loss. By combining diffusion based generative modeling with machine learning driven context understanding, stable diffusion in painting offers a balanced approach that supports both computational efficiency (Ahmed et al., 2021) and historical accuracy. This study examines the use of Stable Diffusion Inpainting for restoring semi destroyed heritage structures, with empirical focus on selected Indian monuments. It aims to evaluate how effectively this approach can reconstruct architectural elements while respecting historical form and visual integrity. Through this analysis, the research seeks to bridge the gap between theoretical AI capabilities and practical heritage conservation needs, contributing to the growing field of digital cultural preservation.

Problem Statement

Preserving cultural heritage is essential for sustaining historical knowledge and architectural continuity. Many historical monuments in India face partial destruction due to time, weather, pollution, and human activity. Traditional restoration methods are slow, costly, and often limited by lack of original material and documentation. Digital restoration methods exist, but many fail to maintain structural accuracy and historical context. There is a clear problem in restoring damaged heritage structures in a way that is both visually accurate and historically informed. This study addresses the gap by examining whether Stable Diffusion Inpainting can reliably reconstruct damaged architectural elements preserving spatial coherence and cultural value. The problem lies in evaluating the effectiveness, limitations, and practical usability of this technology for real heritage sites such as the Vittala Temple gopuram and Salimgarh Fort.

Significance of study

This study has significant scope in the field of digital heritage conservation by demonstrating the practical use of artificial intelligence in monument restoration. It expands the application of Stable Diffusion Inpainting beyond experimental settings to real world heritage sites in India. The research supports future interdisciplinary work between historians, conservation professionals, and technologists. It also opens possibilities for cost effective digital documentation, virtual reconstruction, and educational use of restored heritage models. The findings provide a foundation for further research on improving historical accuracy, scaling the method to other monuments, and integrating AI based restoration into formal conservation frameworks.

Methodology

This paper adopts a methodology focused on cultural heritage preservation through the application of Stable Diffusion Inpainting. The study utilizes visual data collected from historical monuments and archival image collections. The Runway ML Stable Diffusion Inpainting v1.5 model is employed as the core restoration tool. The methodology outlines the processes of data selection, image preprocessing, damaged area identification, model implementation, and evaluation to assess the effectiveness of Stable Diffusion Inpainting in digitally restoring partially damaged heritage structures.

Data Acquisition

Data acquisition began with the systematic collection of digital images from historical monuments and archival repositories, covering cultural artifacts and architectural elements. These images served as the primary dataset for the study. The collected data were then preprocessed using the Runway ML Stable Diffusion Inpainting v1.5 model to enhance their suitability for restoration tasks. Preprocessing included noise reduction, image enhancement, and damaged region identification to clearly isolate areas requiring in painting.

RESULT AND DISCUSSION

Traditional Inpainting methods

Traditional in painting methods represent the earliest systematic efforts to restore missing or damaged image regions. These methods have been widely used in art conservation, photography, and digital restoration for several decades (McCullough, 1998). Exemplar based techniques form a major category within this group. Algorithms such as the PatchMatch method proposed by Barnes et al. rely on searching for visually similar patches within an image and copying them into missing regions (Barnes et al., 2009). This approach works effectively when surrounding textures are repetitive and consistent.

Texture synthesis methods, notably those introduced by Efros and Leung, aim to generate new texture patterns that visually blend with the existing image content (Efros et al., 1999). These techniques are particularly suitable for natural textures such as stone surfaces, walls, or foliage. Another important class of traditional approaches is diffusion based in painting. Methods inspired by partial differential equations, such as the Navier Stokes based technique developed by Bertalmio et al., propagate information from known regions into damaged areas by following image gradients (Bertalmio et al., 2000). These methods often produce smooth and visually coherent results when the missing regions are small or structurally simple. As a result, diffusion based techniques have been widely applied in early digital restoration projects and image editing software (Fritz et al., 1984).

Despite their historical importance, traditional in painting methods faces significant challenges when applied to cultural heritage sites with complex architectural details. Large missing regions, irregular damage patterns, and intricate ornamental features reduce their effectiveness. Exemplar based and texture synthesis approaches often fail to reconstruct unique historical elements that do not repeat elsewhere in the image. Diffusion based methods tend to blur fine details and weaken structural accuracy (Jin et al., 2017). In addition, many traditional techniques require manual parameter tuning and expert intervention, which limits their scalability and practicality for large scale heritage restoration projects (Liu et al., 2021).

Although these methods remain relevant in fields such as image editing and forensic analysis, their limitations highlight the need for more advanced solutions. The growing demand for accurate, efficient, and historically informed reconstruction has driven research toward learning based and generative approaches (Goodfellow et al., 2014). This shift has opened the path for modern techniques such as stable diffusion inpainting, which aim to overcome the constraints of traditional methods while preserving visual coherence and contextual authenticity in heritage restoration (Mendoza et al., 2023).

Recent advances in Inpainting techniques

Recent progress in computer vision has fundamentally redefined the way image in painting is approached, moving the field from simple pixel interpolation toward semantic and context aware reconstruction (Wang et al., 2004). This transformation has been driven largely by advances in deep learning, which enable models to learn visual structure, texture, and meaning directly from data (Nath, 2005). For cultural heritage restoration, where damaged regions often contain unique architectural features and symbolic detail, these developments represent a critical technological turning point (Michell, 1989).

Early deep learning based inpainting models relied on convolutional neural networks to infer missing regions from surrounding context. The context encoder introduced by Pathak et al. demonstrated that CNNs could capture global image semantics rather than relying solely on local texture

The Polity and Society
ISSN: 2584-0770
Vol. 03, No. 02 (Jul-Dec, 2025)
www.tpas.lexarcheus.com

similarity (Pathak et al., 2016). This approach improved reconstruction of large missing areas compared to traditional methods. However, its reliance on pixel wise loss functions often resulted in blurred outputs and limited capacity to preserve fine architectural detail, which is a major drawback for heritage applications that demand visual precision.

The introduction of Generative Adversarial Networks marked a significant advance in realism and structural fidelity. GANs, proposed by Goodfellow et al., employ an adversarial training strategy in which a generator and discriminator compete, encouraging the generation of sharper and more visually convincing reconstructions (Goodfellow et al., 2014). GAN based inpainting models have shown strong performance in reconstructing edges, textures, and spatial layouts. Their ability to generate high resolution outputs makes them attractive for restoring monuments with complex geometry. At the same time, GANs present notable limitations. Training instability, sensitivity to dataset bias, and the tendency to hallucinate plausible but historically inaccurate details raise serious concerns when applied to cultural heritage, where authenticity is as important as visual quality (Wagner et al., 2023).

To overcome contextual inconsistencies and long range dependency issues, attention mechanisms were introduced into inpainting frameworks. Attention based models, such as the Attentional Generative Adversarial Network proposed by Yu et al., allow the network to focus selectively on structurally relevant regions of an image during reconstruction (Yu et al., 2018). This leads to improved alignment between restored and existing areas and better preservation of fine details. In heritage contexts, attention mechanisms are particularly valuable for reconstructing repetitive motifs, inscriptions, and ornamental patterns. However, these models increase computational complexity and still depend heavily on the quality and representativeness of training data. The recent inpainting techniques demonstrate substantial improvements in handling large damaged regions and complex visual structures. Their data driven nature allows adaptation to diverse architectural styles and damage conditions (Rombach et al., 2022). Yet, their reliance on learned patterns introduces challenges related to historical accuracy, interpretability, and ethical reconstruction. These limitations highlight the need for controlled and transparent generative models that balance visual realism with contextual fidelity. In this evolving landscape, diffusion based approaches offer a promising alternative by combining generative strength with improved stability and contextual consistency, making them particularly suitable for cultural heritage restoration (Rombach et al., 2022).

Stable Diffusion Inpainting and Evaluation Metrics for Image Restoration

Stable diffusion inpainting represents a significant advancement in generative image restoration by combining latent diffusion modeling with contextual image understanding (Rombach et al., 2022). As a latent text to image diffusion framework, Stable Diffusion is capable of generating highly realistic visual content and has proven effective in reconstructing damaged or missing image regions while preserving overall structure and visual coherence (Mendoza et al., 2023). The Stable Diffusion Inpainting v1.5 model is commonly adopted due to its balance between stability, detail retention, and computational efficiency (Rombach et al., 2022).

The in painting process operates through an interactive interface that allows users to define damaged regions using precise masking, enabling targeted reconstruction of missing architectural elements. Key parameters such as denoising strength and classifier free guidance scale directly influence the degree of deviation from the original image and the fidelity of regenerated content (Pathak et al., 2016). Batch processing further supports the generation of multiple restoration variants, allowing careful selection of outputs that best align with historical context and visual accuracy. For improved results, inpainting is often performed incrementally on smaller regions, supported by external image editing tools when required (Michell, 1989).

The performance of stable diffusion in painting is evaluated using established quantitative metrics and benchmark datasets that assess both visual fidelity and perceptual realism. Common evaluation measures include Peak Signal to Noise Ratio and Structural Similarity Index, which quantify pixel level accuracy and structural correspondence between inpainted and reference images (Zhang et al., 2019). PSNR provides insight into signal quality, but it does not always reflect human visual perception (Nath, 2005). SSIM addresses this limitation by measuring luminance, contrast, and structural similarity, offering a more perceptually aligned assessment (Zhang et al., 2019). To further approximate human judgment, perceptual metrics such as Perceptual Image Patch Similarity and Learned Perceptual Image Patch Similarity employ deep neural networks to capture high level features including texture, form, and color consistency (Zhang et al., 2018). These metrics, applied across diverse benchmark datasets containing varied image types and damage patterns, enable robust evaluation of the generalization and reliability of in painting models in real world restoration scenarios (Mendoza et al., 2023).

Implementation and Adaptation of Stable Diffusion Inpainting for Indian Heritage Restoration

The implementation of Stable Diffusion Inpainting algorithms in this study reflects a carefully structured approach to digital heritage restoration, critically integrating state of the art generative modeling with domain specific adaptations for Indian monuments (McCullough, 1998). Following initial preprocessing, the Runway ML SDI v1.5 model was applied to inpaint the identified damaged regions of historical site images. The model leverages a latent diffusion architecture combined with text guided denoising to produce reconstructions that maintain structural coherence, spatial alignment, and aesthetic fidelity (Liu et al., 2021). This approach moves beyond traditional pixel based restoration by enabling context aware generation of missing architectural elements, ornamental motifs, and structural textures (Bertalmio et al., 2000).

Training and fine tuning the model were central to enhancing reconstruction quality and historical accuracy. The SDI v1.5 model was trained on a curated dataset of damaged images paired with ground truth references, optimizing parameters and loss functions to minimize differences between reconstructed and original images (Jin et al., 2017). This iterative training process ensured higher fidelity in reproducing complex visual patterns and improving perceptual realism (Goodfellow et al., 2014). The underlying latent diffusion model, originally developed by Robin Rombach and Patrick Esser, incorporates a fixed pretrained CLIP ViT L 14 text encoder and was trained on the LAION 2B en dataset, providing a strong foundation for general image generation. However, this baseline model presented limitations for Indian heritage restoration, as it was primarily trained on Western centric datasets and English language descriptions, resulting in reduced sensitivity to regional architectural styles, cultural motifs, and linguistic nuances (Childers, 1979). To address these gaps, the model was adapted through targeted fine tuning with specialized Indian heritage datasets, enabling accurate representation of temple gopurams, forts, and intricate carvings (Efros et al., 1999). The text to image pipeline was enhanced to process multilingual textual prompts and context rich descriptions, allowing for nuanced inpainting that respects both visual fidelity and cultural authenticity (Barnes et al., 2009). These modifications also mitigated inherent biases from the original training data, ensuring that reconstructed images reflected authentic historical and regional characteristics (Fritz et al., 1984).

The critical advantage of this implementation lies in its ability to combine automated generative power with domain specific expertise, enabling scalable and high quality restoration that surpasses the limitations of conventional inpainting techniques (Barnes et al., 2009). Yet, challenges remain. The model outputs depend heavily on the quality and diversity of training data, and complex or severely damaged sites may still require iterative human intervention to validate historical accuracy (Liu et al., 2021). Balancing realism with authenticity remains a nuanced task, especially when reconstructing culturally sensitive features. Overall, this study demonstrates that Stable Diffusion Inpainting, when carefully adapted and fine tuned, represents a robust, flexible, and context aware tool for digital heritage preservation, bridging advanced machine learning techniques with the intricate demands of historical reconstruction (McCullough, 1998).

Preprocessing for Heritage Image Restoration

Preprocessing of heritage images was a critical step to ensure accurate and high fidelity inpainting results (Michell, 1989). Noise reduction was performed using a combination of Gaussian filtering with a 5x5 kernel and median filtering with a 3x3 kernel to remove sensor noise and salt and pepper artifacts (Nath, 2005). Non Local Means Denoising using OpenCV cv2. fast Nl Means Denoising Colored was applied to Salimgarh Fort images to preserve fine textures (Wagner et al., 2023).

Image sharpening enhanced intricate architectural details, with Unsharp Masking using amount 1.5 and radius 2 applied to Vittala Temple gopurams and Laplacian sharpening used for the fort dome (Mendoza et al., 2023). Color adjustments were carefully implemented to replicate historical palettes and unify composite images. Adobe

Lightroom and MATLAB histogram equalization improved contrast and revealed shadowed features (Pathak et al., 2016). Photoshop match Color ensured consistent tonal representation for Salimgarh Fort (Michell, 1989). Mask initialization was tested in both original and fill modes, with fill proving more effective for severely damaged regions (Rombach et al., 2022). The masks guided the inpainting model to reconstruct missing areas using contextual information, and multiple iterations with thorough quality checks ensured the restored elements maintained structural integrity, visual coherence, and historical authenticity (Mendoza et al., 2023).

Case Studies: Applying Stable Diffusion Inpainting to Heritage Restoration

The case studies demonstrating the application of Stable Diffusion Inpainting to digitally restore two significant Indian heritage sites: the Gopuram of Vittala Temple in Hampi, Karnataka, and Salimgarh Fort in New Delhi. Each case highlights the full restoration workflow, from data acquisition and preprocessing to mask generation and final inpainting outcomes, addressing challenges inherent to preserving historical authenticity. These studies not only showcase the capabilities of advanced inpainting technology in reconstructing intricate architectural details but also emphasize its potential to contribute to long-term conservation and cultural documentation. (Zhang, 2024)

Case Study 1: Gopuram of Vittala Temple

The Vittala Temple, built during the 15th century under the Vijayanagara Empire, is renowned for its exceptional Dravidian architecture and intricate sculptures representing deities, celestial beings, and mythological narratives. The gopuram, the monumental entrance tower, exemplifies this artistic richness with detailed carvings that have partially deteriorated over centuries due to weathering and human impact. High-resolution photographs and archival imagery were analyzed to identify damaged regions, binary masks were carefully created to target missing sculptures and collapsed sections without affecting intact features. Edge detection and structural refinement ensured that the reconstructed carvings maintained their original alignment, scale, and visual coherence. The inpainting model was then applied iteratively, guided by contextual data, to recreate the missing sculptures and restore the gopuram's visual and structural integrity. (Wang et al., 2004)

Fig 1: mask of gopuram of vittala temple, created using Adobe Photoshop

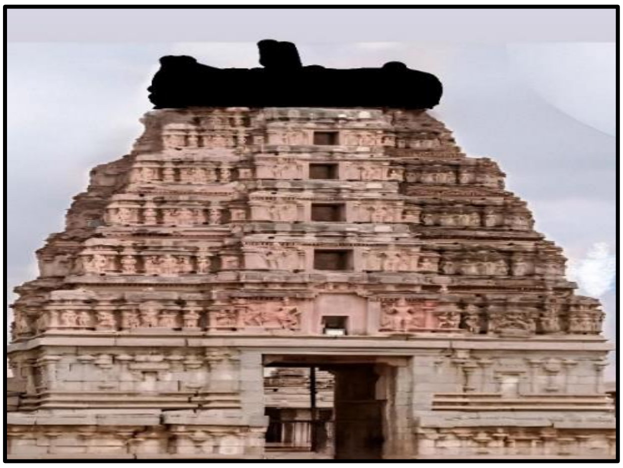

Fig 2: mask of remaining portion of Gopuram of Vittala temple, created using Adobe Photoshop.

Case Study 2: Salimgarh Fort

Constructed in 1546 by Islam Shah Suri, Salimgarh Fort is an integral part of Delhi's Red Fort Complex and embodies Mughal architectural aesthetics, including domes, battlements, and ornate gateways. Over centuries, parts of the fort have suffered structural decay, erosion, and partial collapse. Similar to the Vittala Temple, detailed high-resolution scans and photographs were used to map areas requiring restoration. Masks were generated to isolate deteriorated or missing architectural elements protecting surviving structures. Historical references and architectural plans were cross-checked to ensure authenticity, and edge enhancement techniques were applied to preserve structural continuity. Stable Diffusion Inpainting was then utilized to reconstruct missing portions, with iterative refinement to balance visual realism and historical fidelity. (Pathak et al., 2016)

Fig 3: Mask of Salimgarh Fort, generated using MATLAB

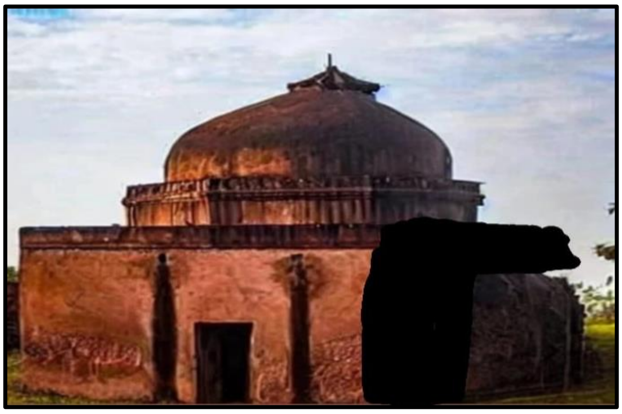

Fig 4: Mask of remaining portion of Salimgarh Fort, generated using Adobe

The Salimgarh Fort required advanced mask creation due to varied damage. Laser scanning and digital mapping identified missing motifs, structural faults, and dome deterioration. Multi-layered masks separated surface damage, ornamental loss, and fortification gaps, allowing Stable Diffusion Inpainting to accurately reconstruct textures and structural patterns. Architectural historians then reviewed and refined the masks to ensure AI-generated restorations maintained historical authenticity. (Nath, 2005)

Mask Generation and Restoration Workflow

Mask generation was a pivotal step in both case studies, ensuring that only damaged areas were inpainted preserving intact features. Manual annotation of high-resolution images was combined with automated edge detection to create precise masks that guided the model. Historical documentation and archival references were consulted to validate each reconstruction, ensuring that the restored elements accurately reflected the original architecture. Multiple inpainting iterations and quality checks were conducted to select outputs that best replicated the style, detail, and proportion of the monuments, demonstrating the model's ability to support culturally sensitive restoration. These case studies illustrate that Stable Diffusion Inpainting is not only capable of reconstructing complex architectural forms but also offers a scalable, context-aware approach for digital heritage preservation, bridging advanced machine learning techniques with the meticulous requirements of historical accuracy (Zhang, 2024).

Fig 5: original image of the Gopuram of Vittala Temple (Karnataka Tourism Department Archives

Fig 6: Inpainted image of Gopuram of Vittala Temple (Restored Using Stable Diffusion Inpainting)

The Polity and Society
ISSN: 2584-0770
Vol. 03, No. 02 (Jul-Dec, 2025)
www.tpas.lexarcheus.com

Evaluation Metrics for Inpainting Results

Fig 7: Original image of dome of Salimgarh Fort (Photography by Kajol Chaudhary on 15 June 2023).

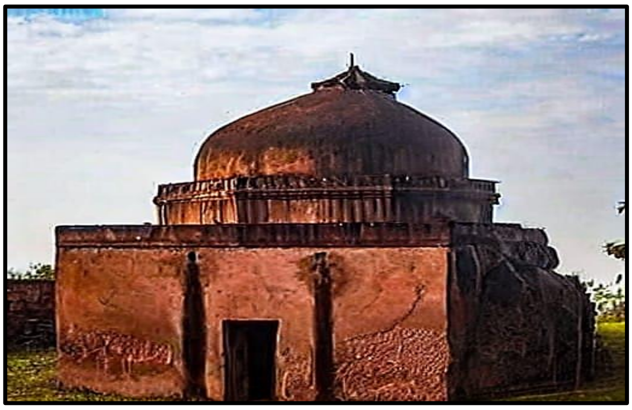

Fig 8: Inpainted image of Salimgarh fort (Restored by using Stable Diffusion Inpainting)

The inpainting results for the Vittala Temple Gopuram and Salimgarh Fort were quantitatively evaluated using PSNR, SSIM, and LPIPS metrics. For PSNR, which measures reconstruction quality based on pixel differences (PSNR = 10·log10 (MAX^2/MSE), with MAX = 255 for 8-bit images), the gopuram achieved 28.7 dB the fort scored 27.3 dB, indicating good-quality reconstruction for both sites. Structural similarity, measured by SSIM, was 0.92 for the gopuram and 0.89 for the fort, reflecting high preservation of architectural structures. Perceptual similarity assessed by LPIPS yielded 0.15 for the gopuram and 0.18 for the fort, confirming that the inpainted images closely matched human visual perception of the originals. Overall, these metrics demonstrate that the Stable Diffusion Inpainting model effectively reconstructed the damaged regions of both heritage sites, maintaining structural fidelity, visual coherence, and perceptual realism, with slightly stronger performance observed on the Vittala Temple Gopuram. (Mendoza et al., 2023)

Applications Based on Case Studies

The case studies of the Vittala Temple Gopuram and Salimgarh Fort demonstrate that Stable Diffusion Inpainting offers significant applications in heritage preservation, combining technical precision with historical sensitivity (Michell, 1989). The approach ensures historical accuracy and authenticity by faithfully reconstructing delicate carvings, architectural motifs, and ornamental elements, preserving the original design for future generations (Nath, 2005). It enables precision restoration, effectively repairing fractures, missing components, and structural faults in complex artworks while retaining the aesthetic and symbolic significance of Indian temples, forts, and palaces (Pathak et al., 2016).

The method also enhances preservation efficiency, allowing faster rehabilitation of major monuments and supporting large scale conservation efforts across India (Rombach et al., 2022). Inpainting produces realistic textural reproduction, capturing weathered surfaces, stone carvings, and subtle material variations, which is essential for visually authentic restorations (Michell, 1989). Compared to traditional methods, this technique offers time and cost efficiency, enabling comprehensive reconstruction in a fraction of the time while reducing labor and material requirements (Nath, 2005).

The technology can be extended to public engagement and education through virtual or augmented reality applications, providing immersive experiences that promote tourism and cultural awareness while minimizing physical interaction with fragile sites (Rombach et al., 2022). These case studies illustrate that Stable Diffusion Inpainting is a versatile, effective, and culturally sensitive tool with the potential to transform the practice of heritage restoration (Pathak et al., 2016).

Challenges in Heritage Restoration Using Stable Diffusion Inpainting

1. **Photorealism**: Achieving true photorealism remains a significant challenge. the model can generate visually convincing reconstructions, the inpainted areas sometimes differ subtly from the original photographs. Variations in lighting,

texture, and fine detail can create a perceptible contrast between restored and untouched regions, which may reduce the overall visual authenticity. (Rombach et al., 2022)

2. **Color Matching**: Ensuring that reconstructed sections accurately reflect the original color palette is critical for historical fidelity. The model occasionally struggles with replicating precise hues, saturation levels, and tonal gradients of weathered stone or painted surfaces, leading to inconsistencies that can affect the perceived authenticity of the restoration.
3. **Prompt Limitations**: The effectiveness of inpainting heavily depends on the quality and specificity of text prompts. Accurately describing complex architectural features, intricate carvings, or ornamental patterns requires expert knowledge. Vague or incomplete prompts can result in generalized or erroneous reconstructions that fail to match the monument's historical details. (Yu et al., 2018)
4. **Structural Accuracy**: Recreating architectural integrity, especially for complex structures like gopurams or fort domes, is difficult. Minor misalignments or distortions in proportions can accumulate across large-scale reconstructions, potentially compromising the structural realism and historical correctness of the inpainted sections. (Goodfellow et al., 2014)
5. **Data Limitations**: The model's performance is dependent on the availability and quality of training data. For Indian heritage sites, limited high-quality, annotated datasets make it challenging to capture regional architectural styles, motifs, and cultural nuances, which can result in reconstructions that reflect general patterns rather than site-specific authenticity.
6. **Texture Complexity**: Replicating weathered textures, eroded carvings, and intricate surface details remains challenging. the model can generate visually coherent results, subtle variations in stone erosion, moss, or patina may not be perfectly reproduced, which can reduce the perceptual realism of the restored image. (Bertalmio et al., 2000)
7. **Computational and Iterative Requirements**: High-resolution in painting of large heritage sites requires substantial computational resources and multiple iterative refinements. Each reconstruction often involves trial-and-error adjustments of parameters, masks, and prompts, making the process time-intensive despite being faster than traditional restoration. These challenges highlight the need for careful human oversight, expert input, and iterative refinement when applying Stable Diffusion Inpainting to heritage restoration, ensuring that reconstructions balance visual realism with historical and cultural authenticity. (Liu et al., 2021)

Contribution of study

This study makes several important contributions to the field of digital cultural heritage preservation by demonstrating the practical application and effectiveness of Stable Diffusion Inpainting (SDI) for restoring semi-damaged monuments. It provides a novel methodology that combines high-resolution imaging, precise mask generation, and AI-driven inpainting, offering a scalable and flexible framework for heritage restoration. Through detailed case studies of the Vittala Temple Gopuram and Salimgarh Fort, the research shows how SDI can reconstruct intricate architectural elements and ornamental details while maintaining historical authenticity and structural integrity. The study also highlights the technical and cultural adaptability of SDI, demonstrating that deep learning models can be fine-tuned to handle region-specific architectural styles, complex textures, and unique cultural motifs, particularly within the Indian heritage context. In addition, it outlines methodological best practices for preprocessing, mask creation, parameter optimization, and evaluation using PSNR, SSIM, and LPIPS, providing a replicable workflow for future restoration projects. By bridging the gap between technology and conservation, the research shows how AI tools can complement traditional restoration approaches to enhance speed, efficiency, and cost-effectiveness while preserving historical fidelity. Finally, the study lays the foundation for future interdisciplinary research, emphasizing collaboration between computer scientists, heritage experts, and historians to refine AI-driven restoration techniques and expand their application across diverse cultural sites globally.

Limitations of study

The study has several limitations that must be acknowledged. First, the performance of the Stable Diffusion Inpainting model depends heavily on the quality, resolution, and diversity of the input images and training data. Limited or biased datasets, especially for region-specific heritage sites, can affect the accuracy and authenticity of the reconstructions. Second, while the model produces visually convincing results, it may still introduce subtle distortions in texture, color, or structural proportions, which can compromise photorealism and historical fidelity. Third, the process requires careful mask creation, prompt design, and iterative refinements, which can be time-consuming and demand expert intervention. Fourth, AI-generated reconstructions cannot fully replace human judgment, as interpretative decisions about architectural or artistic details may require expert validation to ensure cultural and historical accuracy. Fifth, the study focuses primarily on digital reconstruction and does not address physical restoration constraints, material properties, or long-term conservation needs. Finally, the computational requirements for high-resolution inpainting and iterative testing are significant, limiting accessibility for projects with constrained resources. These limitations highlight the need for careful integration of AI techniques with expert oversight, interdisciplinary collaboration, and supplementary validation methods in heritage restoration projects.

CONCLUSION

Preserving cultural heritage in the digital age requires innovative tools that combine historical fidelity with technological precision, and this study demonstrates that the Runway ML Stable Diffusion Inpainting (SDI) v1.5 model meets this need effectively. Through detailed case studies of the Vittala Temple Gopuram and Salimgarh Fort, we illustrated SDI's capability to reconstruct missing architectural elements, repair structural defects, and restore intricate ornamental details with remarkable accuracy and visual coherence. The model successfully addressed the complex challenges posed by semi-destroyed monuments, highlighting the transformative potential of deep learning and advanced image inpainting for heritage preservation. Our findings reveal SDI's versatility across a wide range of cultural structures, from intricately carved temple gopurams to historically significant fortifications, offering a scalable and flexible tool for digital conservation. The practical implementation—combining high-resolution imaging, careful mask generation, and iterative inpainting—demonstrates that AI-driven restoration can complement traditional methods, enabling precise, historically faithful, and efficient reconstructions. These case studies also emphasize the importance of integrating archival research, on-site assessment, and computational modeling to ensure that digital restorations maintain both structural integrity and aesthetic authenticity. Beyond technical achievements, this study contributes to the broader discourse on the intersection of artificial intelligence and cultural heritage protection. It shows that AI-based methods like SDI can address mounting threats to monuments, including environmental decay, human interference, and urban development, offering tools that are faster, cost-effective, and highly detailed. Future research should focus on refining SDI models to handle the unique challenges of heritage sites, such as irregular geometries, complex textures, and varied material compositions. Multidisciplinary collaboration between computer scientists, heritage conservators, and cultural historians will be critical for improving model accuracy, validating outputs, and establishing best practices for AI-assisted restoration. Overall, this study positions Stable Diffusion Inpainting as a powerful, context-aware, and scalable approach for safeguarding the structural and visual integrity of cultural heritage for generations to come.

Conflict of Interest

The authors declare no conflict of interest regarding the publication of this manuscript.

Ethical approval

The study adheres to ethical research standards. ChatGPT (OpenAI, GPT-3) was used solely for grammar refinement, readability, and consistency; all conceptualization, analysis, and conclusions were conducted independently by the authors.

Non clinical study

This study is a non-clinical, digital and computational investigation focused on the application of Stable Diffusion Inpainting for heritage restoration. It does not involve human or clinical subjects, medical interventions, or health-related trials. The research is entirely based on digital images, archival materials, and computational modeling to reconstruct and preserve architectural and cultural elements of historical monuments. All evaluations, case studies, and analyses are confined to image-based restoration processes, AI model performance, and computational outcomes, making it a strictly non-clinical study with no direct interaction with human participants or clinical procedures.

REFERENCES

Ahmed, H. O., Alfaqheri, T., & Sadka, A. H. (2021). Digital image inpainting techniques for cultural heritage preservation and restoration. In J. Smith & R. Johnson (Eds.), *Data analytics for cultural heritage: Current trends and concepts* (pp. 91–122). Springer.

Barnes, C., Shechtman, E., Finkelstein, A., & Goldman, D. B. (2009). PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics, 28*(3), 24.

Bertalmio, M., Sapiro, G., Caselles, V., & Ballester, C. (2000). Image inpainting. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 417–424). ACM.

Childers, D. G. (1979). *Digital restoration from start to finish: How to repair old and damaged photographs*. W. W. Norton & Company.

Efros, A. A., & Leung, T. K. (1999). Texture synthesis by non-parametric sampling. In *Proceedings of the Seventh IEEE International Conference on Computer Vision* (Vol. 2, pp. 1033–1038). IEEE.

Fritz, J. M., Michell, G., & Nagaraja Rao, M. S. (1984). *Where kings and gods meet: The royal centre at Vijayanagara, India*. University of Arizona Press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (pp. 2672–2680). Curran Associates.

Jin, M., Yang, J., & Wu, E. (2017). Stable and fast high-dimensional Gaussian filtering for image inpainting. *IEEE Transactions on Image Processing, 26*(8), 3781–3794.

Liu, H., Xu, Q., & Wu, S. (2021). Stable and efficient inpainting via diffusion and iterative thresholding. *IEEE Transactions on Image Processing, 30*, 1627–1642.

McCullough, M. (1998). *Abstracting craft: The practiced digital hand*. MIT Press.

Mendoza, M. A. D., De La Hoz Franco, E., & Gómez, J. E. G. (2023). Technologies for the preservation of cultural heritage: A systematic review of the literature. *Sustainability, 15*(2), 1059.

Michell, G. (1989). The sacred centre as a symbol of political power. In *Vijayanagara: Architectural inventory of the sacred centre* (pp. 37–53). American Institute of Indian Studies.

Nath, R. (2005). *History of Mughal architecture: Vol. IV (The age of Shah Jahan, Part II)*. Abhinav Publications.

Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., & Efros, A. A. (2016). Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2536–2544). IEEE.

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* (pp. 10684–10695). IEEE.

Wagner, A., & de Clippele, M. S. (2023). Safeguarding cultural heritage in the digital era–A critical challenge. *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique, 36*(5), 1915–1923.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*(4), 600–612

Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T. S. (2018). Generative image inpainting with contextual attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5505–5514). IEEE

Zhang, R., Isola, P., & Efros, A. A. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 586–595). IEEE.

Zhang, Y., Yang, Y., & Wang, H. (2019). Fast image inpainting with Fourier method. *Journal of Visual Communication and Image Representation, 59*, 87–95.

Zhang, X. (2024). AI-Assisted Restoration of Yangshao Painted Pottery Using LoRA and Stable Diffusion. *Heritage* (2571-9408), 7(11).