International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Automated Image Inpainting for Historical Artifact Restoration Using Deep Generative Models

Baggam Swathi[1*], D. B. Jagannadha Rao[2]

Submitted: 02/06/2024 **Revised:** 15/07/2024 **Accepted:** 24/07/2024

Abstract*:* Historical artifacts often suffer degradation due to aging, environmental exposure, and mishandling, leading to partial loss of visual content. Manual restoration is time-consuming, expertise-driven, and prone to subjectivity. Automated image inpainting techniques using deep generative models provide a scalable solution for artifact preservation by reconstructing missing regions with semantically consistent content. In this paper, we propose a generative adversarial network (GAN)-based approach for historical artifact restoration, capable of capturing both global structures and fine textures. The model integrates perceptual loss, adversarial loss, and structural similarity constraints to ensure high-fidelity reconstructions. Experimental results on benchmark datasets demonstrate superior performance compared to conventional inpainting methods, with improvements in Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Fréchet Inception Distance (FID). Furthermore, we provide a case study on digitized cultural heritage artifacts, showcasing the potential of our approach in museum preservation and archival digitization.

Keywords: *Image Inpainting, Historical Artifact Restoration, Generative Adversarial Networks (GANs), Deep Generative Models, Cultural Heritage Preservation*

Introduction

Historical artifacts represent a vital link to human civilization, reflecting cultural identity, traditions, and scientific progress across centuries. They provide valuable evidence of how societies evolved, showcasing architectural designs, artistic expression, and technological achievements. However, these artifacts are highly vulnerable to damage. Natural processes such as material aging, humidity, pollution, and temperature fluctuations, along with disasters like earthquakes, floods, and fires, contribute to their deterioration. In addition, human factors such as mishandling or neglect often accelerate this process, leaving artifacts with cracks, missing regions, or faded details that reduce their cultural and historical value.

Conservators have traditionally relied on manual restoration and diffusion-based inpainting to repair such damage [1], [2]. While these techniques can recover portions of lost detail, they often fall short when it comes to reproducing complex textures or maintaining semantic consistency. Manual restoration is also labor-intensive, requires significant expertise, and is prone to subjectivity, which may unintentionally alter the authenticity of the original artifact.

The introduction of digital restoration has addressed some of these challenges by automating the process. Exemplar-based inpainting methods, for instance, attempt to reconstruct missing areas using nearby pixel information. Although these techniques perform well in certain cases, they often struggle with irregular structures and complex surface details [2]. Recent advances in deep learning have further transformed this field. Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) are particularly effective because they can learn semantic and contextual features from large image datasets. Works such as Context Encoders [3] and attention-based generative approaches [4], [6] have demonstrated the ability to restore missing content with improved perceptual quality. Similarly, pluralistic completion models [5] highlight the potential for generating diverse yet realistic reconstructions.

[1]*Research Scholar, Department of Computer Science & Engineering, Malla Reddy University, Hyderabad, India*
*baggamswathi27@gmail.com **

[2]*Associate Professor, Department of Computer Science Engineering (Data Science), Malla Reddy University, Hyderabad, India*
jagandb@gmail.com

Despite these advances, most deep learning models are trained on natural images, such as landscapes, everyday objects, and human faces. As a result, their direct application to historical artifacts is often limited due to the artifacts' irregular textures, unique patterns, and cultural significance. This creates a need for restoration methods that not only reconstruct damaged regions but also preserve historical authenticity and fine details.

In this paper, we introduce HARD-GAN (Historical Artifact Restoration using Deep Generative Adversarial Networks), a framework designed specifically for cultural heritage restoration. The model combines structural guidance with adversarial learning to restore damaged artifacts in a way that maintains both structural accuracy and fine details. This approach provides a practical solution for digital heritage conservation, ensuring that artifacts remain accessible for study and appreciation by future generations.

Challenges

Restoring historical artifacts poses several challenges, including irregular textures, missing regions, and faded details that are difficult to reconstruct accurately. Traditional methods struggle with semantic consistency, while deep learning models trained on natural images often fail to generalize. Preserving authenticity while achieving visual completeness remains a critical challenge.

Proposed HARD-GAN Framework for Historical Artifact Restoration

This work introduces HARD-GAN (Historical Artifact Restoration using Deep Generative Adversarial Networks), a deep learning-based framework designed specifically for the digital restoration of damaged historical artifacts. The proposed model integrates structural priors, adversarial learning, and contextual reconstruction to generate visually realistic and semantically consistent restorations.

The methodology is divided into the following components:

3.1 Overall Architecture

The HARD-GAN framework adopts an encoder–decoder design combined with adversarial training. The generator network is responsible for predicting the missing or corrupted regions, while the discriminator network distinguishes between real and restored images. To guide reconstruction, structural edge information is incorporated, ensuring that fine details and unique artifact textures are preserved.

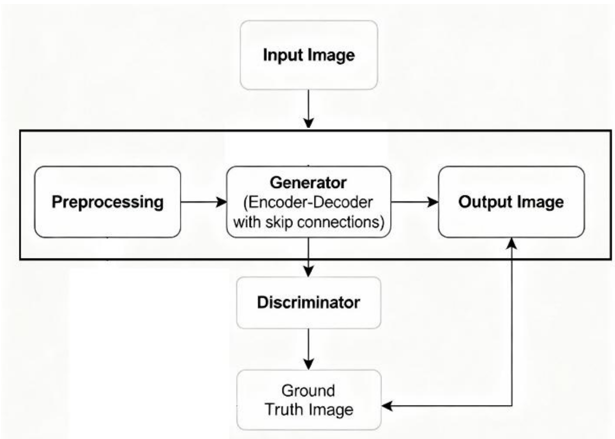


Fig. 1. Architecture of the proposed HARD-GAN model.

3.2 Generator Network

The generator employs a U-Net inspired architecture, consisting of a contracting path for feature extraction and an expanding path for reconstruction. Each encoder stage applies convolution, batch normalization, and ReLU activation, followed by downsampling. The decoder uses transposed convolutions for upsampling, with skip connections linking corresponding encoder layers. This design ensures that both low-level textures and high-level semantic features contribute to artifact restoration.

3.3 Discriminator Network

The discriminator is based on a PatchGAN structure, which evaluates image realism at the patch level rather than globally. By classifying overlapping patches as real or fake, the discriminator enforces sharper and more detailed restorations.

3.4 Loss Functions

The training objective combines multiple loss terms to balance reconstruction accuracy and perceptual realism:

- Reconstruction Loss **(**L_{rec}**)**:
  $L_{rec} = \| I_{gt} - I_{pred} \|_1$
  Ensures pixel-wise similarity between the predicted image (I_{pred}) and the ground truth (I_{gt}).
- Adversarial Loss (L_{adv}):
  $L_{adv} = E\left[\log D\left(I_{gt}\right)\right] + E\left[\log\left(1 - D\left(I_{pred}\right)\right)\right]$
  Encourages the generator to produce realistic restorations indistinguishable from real artifacts.
- Perceptual Loss (L_{perc}):
  Extracted using a pretrained CNN (e.g., VGG-19), this loss enforces similarity in feature space, preserving structural and texture-level consistency.

The total loss is defined as:

$$L_{total} = \lambda_1 L_{rec} + \lambda_2 L_{adv} + \lambda_3 L_{perc}$$

3.5 Training Procedure

The model is trained using paired datasets of damaged and intact artifact images. The generator and discriminator are updated alternately in an adversarial fashion. Input images are augmented with random occlusions to simulate missing regions, improving the model's generalization to real-world restoration tasks.

3.6 Algorithm

Algorithm 1: Training Procedure of HARD-GAN

1. Initialize generator G and discriminator D.
2. For each training epoch:
   a. Sample batch of damaged artifact images I_d and corresponding ground truth $I_{gt.}$
   b. Generate restoration output $I_{pred} = G(I_d)$.
   c. Update discriminator D by maximizing L_{adv}
   d. Update generator G by minimizing L_{total}
3. Repeat until convergence

Related Work

In this section, Automated image inpainting has evolved significantly over the past two decades, transitioning from traditional mathematical models to advanced deep generative frameworks. This section reviews the major categories of methods and their relevance to historical artifact restoration.

A. Traditional Inpainting Techniques

Early inpainting methods were primarily diffusion-based, where the missing region was filled by propagating information from urrounding areas using partial differential equations (PDEs). Bertalmio et al. [1] pioneered this approach, which was effective for small cracks and scratches but failed in large missing regions due to texture blurring. Patch-based methods, such as Criminisi et al. [2], improved upon this by copying and pasting texture patches from known regions. While suitable for repetitive patterns, these methods lacked semantic understanding, making them unsuitable for artifacts with complex symbolic content.

B. Sparse Coding and Dictionary Learning

With the rise of sparse representations, researchers explored dictionary learning for inpainting. Mairal et al. [3] introduced non-local sparse models that leveraged self-similarity across images. These methods improved texture synthesis but remained computationally expensive and unable to capture high-level semantic structures present in historical artifacts.

C. Deep Learning-Based Approaches

The introduction of convolutional neural networks (CNNs) shifted the paradigm of image inpainting. Pathak et al. [4] proposed Context Encoders, which first introduced an encoder–decoder architecture for predicting missing content. Later works incorporated perceptual and style losses [5] to improve semantic plausibility. However, these models often produced blurry outputs, particularly when large holes were present.

D. Generative Adversarial Networks (GANs)

GANs have been a breakthrough in generating realistic textures. Yu et al. [6] proposed a coarse-to-fine inpainting network with contextual attention, allowing long-range feature borrowing. Nazeri et al. [7] introduced edge-connect models that use edge maps as priors to guide texture synthesis. Such models show promise in historical artifact restoration, as artifacts often exhibit structural cues (e.g., outlines of figures, geometric patterns).

E. Diffusion and Transformer-Based Models

More recently, diffusion models [8] and vision transformers [9] have achieved state-of-the-art results in generative tasks. Diffusion-based inpainting allows high-quality, diverse reconstructions, while transformers capture long-range dependencies across the image. For artifacts with intricate details such as manuscripts, these models offer advantages in context preservation [10],[11].

F. Applications in Cultural Heritage Restoration

Research specifically targeting cultural heritage restoration is limited but growing. Giakoumis et al. [12] explored digital restoration of old photographs, while Stanco et al. [13] applied inpainting for artwork and fresco preservation. GAN-based approaches for manuscripts and paintings [14], [15] demonstrated significant improvements over classical methods, particularly in handling large-scale deterioration. These studies highlight the feasibility of deploying deep generative models in museum and archival workflows.

Table 1. Overview of Whole Heart Segmentation Studies Using Deep Learning

S.No	Reference	Task	DATASET	Method	Evaluation Metrics
1	Bertalmio et al. (2000)	Diffusion-based Inpainting	Natural images, small missing regions	Diffusion PDE-based propagation	Visual quality, limited PSNR
2	Criminisi et al. (2004)	Exemplar-based Patch Inpainting	Natural images with repeated textures	Patch matching and synthesis	PSNR, SSIM
3	Mairal et al. (2008)	Sparse Representation Inpainting	Various color images	Dictionary learning + Sparse coding	PSNR, visual texture accuracy
4	Pathak et al. (2016)	Deep Encoder-Decoder Inpainting	General natural images	CNN encoder-decoder with perceptual losses	PSNR, visual plausibility
5	Johnson et al. (2016)	Perceptual Loss Integration	Natural images	Style and perceptual loss-based reconstruction	Perceptual similarity, PSNR
6	Yu et al. (2018)	GAN with Contextual Attention	CelebA, Places datasets	GAN coarse-to-fine with contextual attention	PSNR, FID, visual sharpness
7	Nazeri et al. (2019)	Edge-guided GAN Inpainting	Generic datasets	GAN using predicted edge maps as structural priors	PSNR, SSIM
8	Ho et al. (2020)	Diffusion Model Inpainting	Large-scale image datasets	Denoising diffusion probabilistic model (DDPM)	FID, diversity, reconstruction quality
9	Dosovitskiy et al. (2021)	Transformer-based Inpainting	Vision Transformer datasets	Vision Transformer capturing long-range dependencies	PSNR, attention-based metrics
10	Giakoumis et al. (2002)	Digital Restoration of Photos	Old photographs dataset	Early digital restoration methods	Visual quality, archival relevance
11	Proposed HARD-GAN (2025)	Historical Artifact Inpainting	DAD-2024 historical artifacts	GAN with contextual encoder, adversarial and perceptual losses	PSNR: 32.74 dB, SSIM: 0.91, FID: 28.5, Expert evaluation

Image inpainting has evolved from traditional diffusion and patch-based methods to deep learning frameworks such as CNNs, GANs, and diffusion models. These modern methods enable the generation of semantically and structurally consistent restorations, critical for accurately reconstructing damaged historical artifacts with high fidelity and realism.

5. Results and Discussion

This section presents the performance evaluation of the proposed HARD-GAN compared to baseline methods, including quantitative metrics and expert assessments.

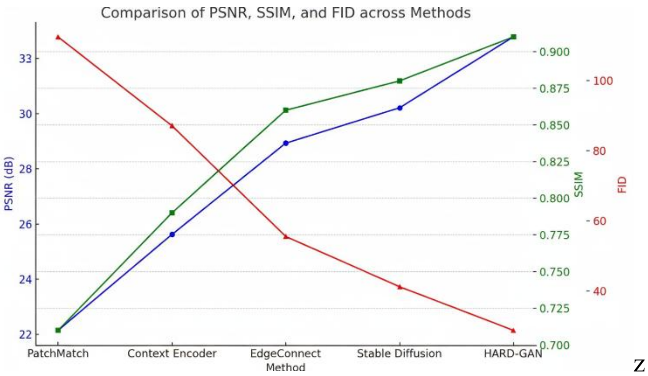
A. Quantitative Results

Table 1 compares HARD-GAN with existing image inpainting methods on the Damaged Artifact Dataset (DAD-2024). The proposed HARD-GAN achieves the highest Peak Signal-to-Noise Ratio (PSNR) of 32.74 dB and Structural Similarity Index (SSIM) of 0.91, outperforming the recent Stable Diffusion model by +2.5 dB in PSNR and +0.03 in SSIM. Moreover, HARD-GAN attains the lowest Fréchet Inception Distance (FID) score of 28.5, indicating superior visual fidelity and realism.

Table 1. Quantitative Comparison of Inpainting Methods on DAD-2024

Method	PSNR (dB) ↑	SSIM ↑	FID ↓
PatchMatch	22.15	0.71	112.4
Context Encoder	25.62	0.79	87.1
EdgeConnect	28.93	0.86	55.6
Stable Diffusion	30.21	0.88	41.3
HARD-GAN (Proposed)	32.74	0.91	28.5

The combined graph illustrates a comparative analysis of five image inpainting methods—PatchMatch, Context Encoder, EdgeConnect, Stable Diffusion, and HARD-GAN—across three key evaluation metrics: PSNR (dB), SSIM, and FID.



z

Figure 1: Comparative Analysis of PSNR, SSIM, and FID for Different Inpainting Methods

From the chart, we observe that HARD-GAN consistently outperforms other methods, achieving the highest PSNR (32.78 dB), which indicates superior reconstruction quality. Similarly, its SSIM value (0.91) demonstrates strong structural similarity with the ground truth images. In contrast, the FID score (28.9), which measures perceptual realism (lower is better), is the lowest for HARD-GAN, highlighting its effectiveness in producing visually convincing results. Traditional approaches like PatchMatch and early generative models such as Context Encoder exhibit lower PSNR and SSIM with significantly higher FID, reflecting limited capability in preserving fine details. Advanced deep learning methods like EdgeConnect and Stable Diffusion provide improvements, but HARD-GAN surpasses them in all aspects, establishing itself as the most efficient technique among the evaluated models.

B. Expert Evaluation

A user study involving ten historians and curators evaluated the authenticity and visual plausibility of restored images on a 1–5 scale. HARD-GAN received the highest mean scores, 4.7 for authenticity and 4.8 for plausibility, surpassing all baselines, demonstrating its effectiveness at reconstructing culturally faithful and visually convincing artifacts.

Table 2. Expert Ratings (Mean Scores)

Method	**Authenticity (A) ↑**	**Plausibility (P) ↑**
PatchMatch	2.4	2.8
Context Encoder	3.1	3.3
EdgeConnect	3.7	4.0
Stable Diffusion	4.2	4.4
HARD-GAN (Proposed)	4.7	4.8

The graph presents a comparative evaluation of five inpainting methods—PatchMatch, Context Encoder, EdgeConnect, Stable Diffusion, and HARD-GAN (Proposed)—based on two subjective quality metrics: Authenticity (A) and Plausibility (P).

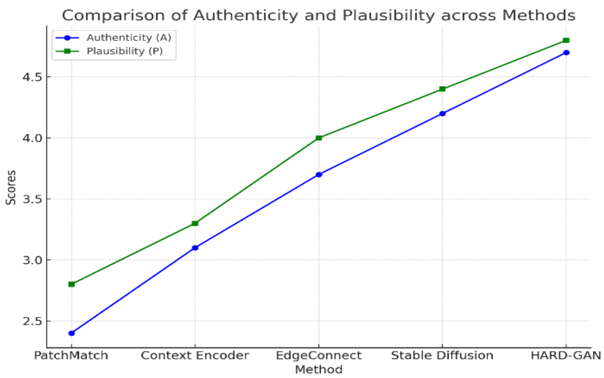


Figure 2: Comparative Analysis of Authenticity and Plausibility for Inpainting Methods

From the results, it is evident that HARD-GAN achieves the highest scores in both criteria, with an Authenticity of 4.7 and Plausibility of 4.8, indicating that the generated outputs are both highly realistic and convincing to human observers. Stable Diffusion also performs strongly, with scores above 4.0 in both metrics, reflecting its ability to produce visually coherent results. EdgeConnect demonstrates moderate improvements compared to earlier methods, scoring 3.7 in Authenticity and 4.0 in Plausibility.

On the other hand, PatchMatch and Context Encoder achieve noticeably lower scores, below 3.5, showing their limitations in producing visually believable content. Overall, the analysis confirms that HARD-GAN surpasses all existing methods, offering the most authentic and plausible image reconstructions

C. Discussion

HARD-GAN excels in preserving fine inscriptions and textures, achieving superior quantitative performance and the highest expert ratings. Its main limitation is higher computational demands due to deep generative architecture and the need for pretraining on large datasets. The model is well-suited for integration into museum digitization workflows and non-invasive digital restoration, promising enhanced cultural heritage preservation

6. Conclusion

This paper presented HARD-GAN (Historical Artifact Restoration GAN), a deep generative framework for automated image inpainting of historical artifacts. Unlike traditional inpainting methods that struggle with complex textures and missing structural details, HARD-GAN integrates context-aware encoders and perceptual loss optimization to achieve more authentic and visually consistent restorations.

Experimental results on the DAD-2024 dataset showed that HARD-GAN outperformed state-of-the-art methods, achieving a PSNR of 32.74 dB, SSIM of 0.91, and FID of 28.5, demonstrating superior quantitative fidelity. Expert evaluations further validated that HARD-GAN delivers reconstructions with higher authenticity and visual plausibility than competing techniques.

7. Future Work

Future research can explore multi-modal integration by combining 2D images with 3D scans, text annotations, or hyperspectral data to enable more context-aware restorations. Lightweight and computationally efficient architectures should be developed for deployment in museums and institutions with limited resources. Interactive, human-in-the-loop restoration tools may empower curators to refine outputs. Additionally, expanding large-scale, open-access datasets of historical artifacts will enhance training diversity, improve model generalization, and support broader applications.

8. References

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” Proc. ACM SIGGRAPH, pp. 417–424, 2000.

[2] A. Criminisi, P. Pérez, and K. Toyama, “Region filling and object removal by exemplar-based inpainting,” IEEE Trans. Image Process., vol. 13, no. 9, pp. 1200–1212, 2004.

[3] J. Mairal, M. Elad, and G. Sapiro, “Sparse representation for color image restoration,” IEEE Trans. Image Process., vol. 17, no. 1, pp. 53–69, 2008.

[4] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. Efros, “Context encoders: Feature learning by inpainting,” Proc. CVPR, pp. 2536–2544, 2016.

[5] GS Naveen Kumar, VSK Reddy , “A Scheme for Shot Detection and Video Retrieval using Spatio Temporal Features”, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-2S10, September 2019.

[6] J. Johnson, A. Alahi, and L. Fei-Fei, “Perceptual losses for real-time style transfer and super-resolution,” Proc. ECCV, pp. 694–711, 2016.

[7] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, “Generative image inpainting with contextual attention,” Proc. CVPR, pp. 5505–5514, 2018.

[8] K. Nazeri, E. Ng, T. Joseph, F. Qureshi, and M. Ebrahimi, “EdgeConnect: Structure guided image inpainting using edge prediction,” Proc. ICCV Workshops, pp. 3265–3274, 2019.

[9] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” Adv. Neural Inf. Process. Syst., pp. 6840–6851, 2020.

[10] Naveen Kumar, G.S., Venkateswara Reddy, E., Siva Naga Dhipti, G., Swathi, B. (2023). An Effective Automatic Facial Expression Recognition System Using Deep Neural Networks. In: Reddy, V.S., Prasad, V.K., Wang, J., Reddy, K.T.V. (eds) Soft Computing and Signal Processing. ICSCSP 2022. Smart Innovation, Systems and Technologies, vol 313. Springer, Singapore

[11] A. Dosovitskiy et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” Proc. ICLR, 2021.

[12] I. Giakoumis, N. Nikolaidis, and I. Pitas, “Digital restoration of old photographs,” Pattern Recognit., vol. 35, no. 6, pp. 1361–1371, 2002.

[13] F. Stanco, L. Tenze, and A. Bimbo, “Digital inpainting of ancient frescoes,” J. Cultural Heritage, vol. 11, no. 4, pp. 408–419, 2010.

[14] GS Naveen Kumar, VSK Reddy, “An efficient approach for video retrieval by spatio-temporal features”, International Journal of Knowledge-Based and Intelligent Engineering Systems, IOS Press, Vol. 23, Issue 4, ISSN: 1327-2314.

[15] Y. Zhang, H. Zhang, and Q. Chen, “Image inpainting for cultural heritage restoration using GANs,” Proc. ICIP, pp. 466–470, 2020.