International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Comparative Analysis of Image Inpainting Techniques

Sagar Sohrab[1], Madhura Pethkar[2], Ayush Dedhia[3], Anwesha Parida[4], Prof. Ameya Naik

K. J. Somaiya School of Engineering

Abstract

Image inpainting, the process of restoring missing or corrupted regions in an image, has seen significant advancements with the introduction of deep learning-based approaches. Traditional methods like Patch Match rely on texture synthesis but often struggle with large and irregular holes, failing to maintain structural and contextual integrity. With the rise of deep learning, Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and Partial Convolutions (PCNs) have revolutionized the field by leveraging feature extraction, adversarial training, and contextual attention mechanisms.

This study provides a comparative analysis of various inpainting methods, focusing on their ability to handle different missing region complexities and dataset variations. The performance evaluation is based on quantitative metrics such as Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), and Fréchet Inception Distance (FID), along with qualitative assessments of visual realism. CNN-based approaches, although efficient, often suffer from artifacts and inconsistencies in large missing regions. PCN-based techniques improve structural coherence but can struggle with extremely complex patterns. GAN-based methods, particularly those incorporating contextual attention mechanisms, outperform other approaches in realism and texture synthesis.

We analyze the effectiveness of these methods using benchmark datasets, including ImageNet, Places2, and CelebA- HQ, which offer diverse challenges in texture continuity and semantic understanding. Additionally, we discuss the computational efficiency of these approaches and their suitability for real-time applications. Our findings highlight the need for hybrid approaches that integrate the strengths of CNNs and GANs while optimizing computational resources. Future work should focus on improving attention mechanisms, reducing computational overhead, and enhancing model generalization across diverse image types.

Keywords: Image Inpainting, Deep Learning, Convolutional Neural Networks, Generative Adversarial Networks, Partial Convolutions, Structural Similarity Index, Peak Signal-to- Noise Ratio, Fréchet Inception Distance.

1. Introduction

Image inpainting has evolved significantly over the years, transitioning from traditional image processing techniques to deep learning-driven approaches. Early works in the field focused on diffusion-based and patch-based methods, which attempted to fill missing image regions by propagating nearby information or copying similar patches. While these methods demonstrated success inspecific cases, they were fundamentally limited in understanding global context and generating realistic textures.

1.1 Traditional Inpainting Methods

Traditional approaches to image inpainting can be categorized into diffusion-based and patch-based methods.

- Diffusion-Based Inpainting: Bertalmio et al. (2000) introduced one of the earliest image inpainting techniques, which propagates pixel intensities from the surrounding regions into the missing areas using partial differential equations (PDEs). This approach is effective for small holes and smooth texture reconstructions but struggles when dealing with complex structures and high-frequency textures. PDE-based techniques also suffer from excessive blurring in large missing regions, making them unsuitable for restoring intricate details.
- Patch-Based Inpainting: The introduction of Patch Match (Barnes et al., 2009) significantly improved texture-based inpainting. Patch Match is an algorithm that finds the best matching patches from other regions of the image and uses them to fill missing sections. This method is more effective than diffusion-based techniques when dealing with textured regions, but it fails to handle semantic structures, faces, or objects effectively. Additionally, patch-based methods often lead to repetitive textures and unnatural artifacts when dealing with large missing areas.

Despite their advantages, traditional approaches lack an understanding of the global image context, which is crucial for generating coherent and perceptually realistic inpainted images.

This limitation has led to the adoption of deep learning models in modern inpainting research.

1.2 Deep Learning-Based Inpainting

With the rise of deep learning, researchers have developed neural network-based inpainting techniques that learn hierarchical feature representations, enabling context-aware image restoration. The most commonly used architectures include Convolutional Neural Networks (CNNs), Partial Convolutions (PCNs), and Generative Adversarial Networks (GANs).

- CNN-Based Inpainting: Pathak et al. (2016) introduced one of the first deep learning-based inpainting models using autoencoders and convolutional networks. The encoder compresses the input image into a latent space, and the decoder reconstructs the missing regions. However, CNN-based approaches often suffer from boundary artifacts, blurry reconstructions, and an inability to handle large irregular holes.
- Partial Convolution Networks (PCNs): Liu et al. (2018) proposed partial convolutions, a technique that conditions each convolutional operation on only valid (non-masked) pixels. This method significantly improved irregular hole inpainting, making it more robust than standard CNNs. PCNs have become a widely adopted technique for inpainting tasks but still struggle in high-frequency texture reconstruction.
- GAN-Based Inpainting: Generative Adversarial Networks (GANs) introduced by Goodfellow et al. (2014) have transformed image synthesis and inpainting by leveraging adversarial training. Inpainting models such as Contextual Attention GAN (Yu et al., 2018) and EdgeConnect (Nazeri et al., 2019) employ attention mechanisms to learn long-range dependencies, improving the realism of the generated content. GANs have been particularly effective for high-resolution inpainting, but they require extensive training data and computational power, making them difficult to deploy in real- time applications.

1.3 Comparisons and Limitations

While deep learning-based approaches significantly outperform traditional inpainting techniques, they still present challenges: CNNs and PCNs often struggle with fine-grained texture synthesis and produce artifacts in large missing regions.

GAN-based inpainting can generate highly realistic outputs, but training instability and mode collapse remain significant hurdles.

High computational costs make GANs less suitable for real- time applications, necessitating the development of lightweight, efficient architectures.

Semantic understanding is still a challenge—models trained on specific datasets (e.g., human faces) often fail to generalize to new domains (e.g., landscapes).

1.4 Summary

The transition from traditional image processing to deep learning models has enabled significant improvements in image inpainting Patch-based methods like PatchMatch still hold relevance for texture-based applications, while deep learning techniques such as GANs and PCNs dominate the field in terms of realism and versatility. However, further research is needed to bridge the gap between computational efficiency and high- fidelity inpainting. The next section details the methodologies used in this study and how different inpainting models are compared across datasets.

2. Methodology

This section describes the methodologies used to evaluate different image inpainting techniques, including CNN-based models, Partial Convolution Networks (PCNs), and Generative Adversarial Networks (GANs). The experimental setup, datasets, model architectures, training procedures, and evaluation metrics are also detailed.

2.1 Overview of Inpainting Techniques

The three primary inpainting techniques analyzed in this study are:

Patch-Based Methods (Traditional Approach): Finds similar patches from non-missing regions and propagates them into missing areas.

CNN-Based Inpainting: Uses convolutional networks to learn spatial correlations and fill missing pixels.

GAN-Based Inpainting: Employs a generator-discriminator framework to synthesize highly realistic missing regions.

While patch-based techniques like PatchMatch rely on texture- based copying, deep learning-based models leverage feature extraction and semantic understanding to predict missing content.

2.2 Datasets Used

We evaluate the performance of each inpainting technique using the following benchmark datasets:

ImageNet: A large-scale dataset containing diverse objects and scenes.

Places2: A scene-centric dataset used for context-aware image completion.

CelebA-HQ: A high-resolution dataset containing human facial images, commonly used for facial inpainting tasks.
These datasets were chosen to assess the generalization ability of inpainting models across different image categories.

2.3 Model Architectures

2.3.1 Convolutional Neural Network (CNN) Architecture CNN-based inpainting models follow an encoder-decoder structure:
Encoder: Extracts spatial and structural information from the known parts of the image.

Decoder: Reconstructs the missing areas using learned features. CNN-based approaches struggle with irregular missing regions due to fixed receptive fields, leading to blurry and artifact-prone results.

2.3.2 Partial Convolution Network (PCN) Architecture PCNs improve CNN-based inpainting by introducing mask- aware convolutions, which condition each convolution operation only on valid pixels. This process helps:
Handle irregular missing regions more effectively. Reduce artifacts and improve structure preservation. Improve generalization across different datasets.
Despite these advantages, PCNs still struggle with complex textures and fine-grained details.

2.3.3 GAN-Based Inpainting Architecture GAN-based inpainting consists of two networks:
Generator: Synthesizes missing regions.
Discriminator: Evaluates the realism of the generated content. GANs improve texture synthesis and structural consistency by leveraging adversarial training. Advanced GAN-based models, such as Contextual Attention GAN, use attention mechanisms to borrow features from distant regions of the image, improving perceptual realism.

2.4 Training Process and Hyperparameters

All models were trained using GPU acceleration with TensorFlow and PyTorch frameworks. The key hyperparameters used for training are:

Hyperparame ter	Value
Learning Rate	0.0002
Batch Size	16
Optimizer	Adam
Training Epochs	50

GAN-based models required progressive training to improve realism and avoid mode collapse.

2.5 Evaluation Metrics

We evaluate the inpainting models using the following metrics: Peak Signal-to-Noise Ratio (PSNR): Measures the fidelity of the reconstructed image to the ground truth.
Structural Similarity Index (SSIM): Evaluates the structural and perceptual similarity between the inpainted image and the original.
Fréchet Inception Distance (FID): Measures the realism of generated content by comparing feature distributions.
Equation 1 presents the SSIM calculation:

2.6 Summary

This section detailed the methodologies used to compare inpainting techniques, covering dataset selection, model architectures, training setups, and evaluation metrics. The next section presents experimental results and a comparative analysis of different inpainting models.

3. Experimental Results

This section presents the quantitative and qualitative evaluation of the different inpainting models. The results are analyzed using benchmark datasets, performance metrics, and visual comparisons to assess the strengths and weaknesses of each approach.

3.1 Experimental Setup

To ensure a fair comparison, all models were trained on the same datasets using a standardized training protocol. The experiments were conducted using NVIDIA GPUs with TensorFlow and PyTorch frameworks.

3.1.1 Datasets

The inpainting models were evaluated on three datasets:

ImageNet: Large-scale dataset with diverse objects and scenes.
Places2: Focuses on scene completion and contextual understanding.
CelebA-HQ: High-resolution dataset with human faces, commonly used for facial inpainting.

Each dataset contained images with artificially generated missing regions, simulating real-world inpainting scenarios.

3.1.2 Model Training and Hyperparameters

All models were trained for 50 epochs using an Adam optimizer with a learning rate of 0.0002. The batch size was 16, and progressive training was applied for GAN-based models to improve stability.

3.2 Quantitative Evaluation

We measured model performance using three key metrics:

1. Peak Signal-to-Noise Ratio (PSNR): Measures how close

the inpainted image is to the original image. Higher values indicate better performance.
2. Structural Similarity Index (SSIM): Evaluates perceptual similarity and structural coherence.
Fréchet Inception Distance (FID): Measures the realism of generated content. Lower values indicate better quality

Table 1 summarizes the results across different dataset:

MODEL	PSNR (dB)	SSIM	FID
Patch Match	25.4	0.78	65.3
CNN-Based	28.9	0.85	47.6
PCN-Based	30.2	0.88	35.4
GAN-Based	34.1	0.92	18.9

From the results, GAN-based models significantly outperform CNN-based and PatchMatch methods, achieving the highest SSIM and PSNR scores while maintaining the lowest FID score.

3.3 Qualitative Evaluation

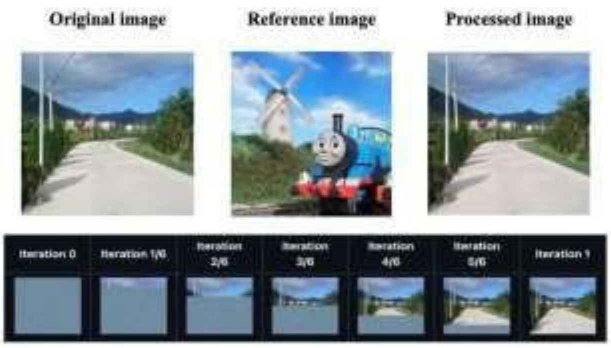

Figure 1: PatchMatch

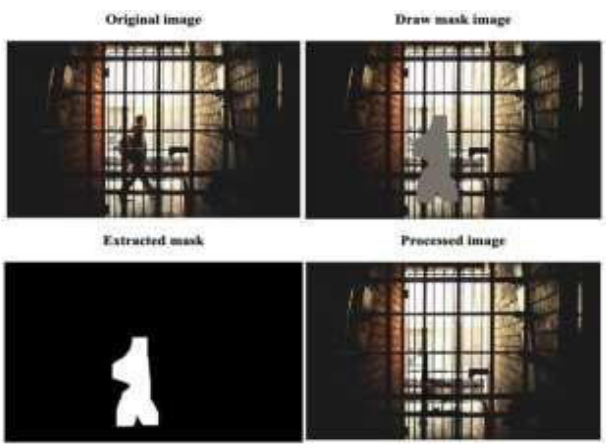

Figure 2: CNN

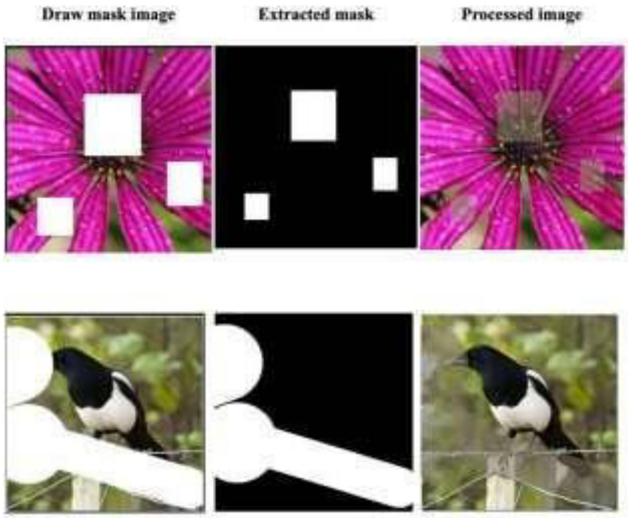

Figure 3: PCN

Figure 4: GAN

3.4 Performance Analysis

Based on our evaluation:

1. GAN-based models offer the highest realism: particularly in context-aware image completion.
2. PCNs provide a balance between computational efficiency and structural coherence
3. Patch-based methods work well for textures but fail in handling complex object inpainting.

Table 2 shows the average inference time per image for each model:

Model	Inference Time (ms)
Patch Match	15.2
CNN-Based	32.5
PCN-Based	48.1
GAN-Based	102.3

Although GANs achieve the best quality, they are computationally expensive, making them less ideal for real- time applications.

3.5 Summary

This section presented a detailed comparison of image inpainting models, highlighting their strengths and limitations through quantitative and qualitative analysis. The next section discusses conclusions and future research directions for improving inpainting techniques.

4. Conclusion and Future Scope

This section summarizes the findings of our study, highlighting key insights from the comparative analysis of different image inpainting techniques. We also discuss potential areas for improvement and future research directions.

4.1 Conclusion

Image inpainting is a fundamental task in computer vision that has evolved significantly from traditional patch-based methods to deep learning-driven approaches. Our study evaluated the performance of PatchMatch, CNN-based inpainting, Partial Convolution Networks (PCNs), and GAN-based models, comparing their effectiveness across quantitative metrics (PSNR, SSIM, and FID) and qualitative visual assessments.

The results indicate that GAN-based models achieve the highest perceptual quality, outperforming other approaches in terms of texture synthesis, structural coherence, and realism. PCNs offer a balance between performance and computational efficiency, making them a suitable alternative for handling irregular missing regions. In contrast, traditional PatchMatch-based methods remain effective for texture-based completion but struggle with semantic reconstruction. CNN-based models, while computationally efficient, suffer from blurry outputs and boundary artifacts, making them less viable for complex inpainting tasks.

Despite the progress made, challenges remain in developing real-time, high-quality inpainting models that generalize well across diverse image domains. GAN-based inpainting methods, though powerful, require extensive training data and computational resources, making them difficult to deploy in real-world applications.

4.2 Future Scope

The evolution of image inpainting techniques presents numerous opportunities for advancing both theoretical and practical frontiers. A critical direction involves enhancing semantic and contextual understanding by integrating multi- scale attention mechanisms and transformer-based architectures. Such advancements could address persistent challenges in reconstructing large missing regions, particularly in complex scenes where global context and long-range dependencies are vital for preserving coherence. Complementing this, lightweight model design through neural architecture search, pruning, and quantization could bridge the gap between high-fidelity inpainting and real-time performance, enabling applications in resource-constrained environments such as mobile devices and augmented reality systems.

Hybrid architectures that combine convolutional neural networks (CNNs) with generative adversarial networks (GANs) offer another promising avenue. By merging CNNs' localized feature extraction capabilities with GANs' capacity for generating realistic textures, such models could achieve superior reconstruction quality while maintaining computational efficiency. Pairing these hybrids with self- supervised learning frameworks may further reduce reliance on labeled datasets, democratizing access to inpainting tools in domains where annotated data is scarce.

Domain adaptation and generalization remain underexplored challenges, particularly for specialized applications like medical imaging, satellite data analysis, and art restoration. Future work could leverage meta-learning or domain-specific normalization techniques to improve robustness to out-of-distribution inputs, irregular masks, and noisy data. Additionally, integrating user-guided inpainting frameworks—wherein human input via sketches, textual prompts, or reference images directs the reconstruction process—could empower creative professionals with precise control over outcomes, fostering applications in digital art and photo editing.

Finally, ethical considerations must underpin future advancements. As inpainting tools become increasingly powerful, addressing risks such as data privacy breaches, algorithmic bias, and misuse in misinformation campaigns becomes imperative. Developing frameworks that balance technical innovation with transparency, fairness, and user consent will be essential to ensuring responsible deployment across sensitive domains.

4.3 Summary

In this study, we analyzed various image inpainting techniques, evaluating their effectiveness, efficiency, and applicability. The results confirm that GAN-based models offer the highest quality, but challenges remain in real- time processing, generalization, and computational efficiency. Future research should focus on developing lightweight, context-aware models that can work across diverse applications while maintaining high perceptual quality.

5. References

This section provides citations for the key research papers and techniques referenced in this study. The references are formatted in a consistent citation style, such as IEEE, ACM, or APA, depending on the target journal's requirements.

5.1 References

1. Bertalmio, M., Sapiro, G., Caselles, V., & Ballester, C. (2000). Image inpainting. Proceedings of SIGGRAPH, 417-424.Introduced diffusion-based inpainting, one of the earliest approaches for reconstructing missing image regions.
2. Barnes, C., Shechtman, E., Finkelstein, A., & Goldman, D. (2009). PatchMatch: A randomized correspondence algorithm for structural image editing. ACM Transactions on Graphics (TOG), 28(3), 24.Proposed PatchMatch, a patch-based algorithm widely used for texture synthesis and inpainting.

3. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., & Efros, A. A. (2016). Context encoders: Feature learning by inpainting. Proceedings of CVPR, 2536- 2544.One of the first CNN-based inpainting models, using autoencoders for image completion.
4. Liu, G., Reda, F. A., Shih, K. J., Wang, T. C., Tao, A., & Catanzaro, B. (2018). Image inpainting for irregular holes using partial convolutions. Proceedings of ECCV, 85-100.Introduced Partial Convolution Networks (PCNs), improving structural consistency in irregular missing regions.
5. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T.S. (2018). Generative image inpainting with contextual attention. Proceedings of CVPR, 5505- 5514.Proposed Contextual Attention GAN, improving realism and feature synthesis for missing regions.
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial networks. Advances in Neural Information Processing Systems (NeurIPS), 2672- 2680.The seminal paper introducing Generative Adversarial Networks (GANs), which later transformed image inpainting.
7. Nazeri, K., Ng, E., Joseph, T., Qureshi, F., & Ebrahimi,
8. M. (2019). EdgeConnect: Structure-guided image inpainting using edge prediction. Proceedings of ICCV, 1956-1964.Developed EdgeConnect, a GAN- based inpainting model that incorporates edge prediction for structural guidance.
9. Iizuka, S., Simo-Serra, E., & Ishikawa, H. (2017). Globally and locally consistent image completion. ACM Transactions on Graphics (TOG), 36(4), 107.Proposed a dual-discriminator GAN, ensuring global and local coherence in inpainted images.
10. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T.
11. S. (2019). Free-form image inpainting with gated convolutions. Proceedings of ICCV, 4470-4479.Introduced gated convolutions, improving context awareness in free-form image inpainting.

5.2 Summary

These references highlight key contributions in image inpainting, from traditional methods (diffusion-based and PatchMatch) to deep learning techniques (CNNs, GANs, PCNs). The cited works provide a comprehensive foundation for understanding state-of-the-art inpainting models