A Review of Modern Image Steganography Techniques

[1]Rajneesh Choudhary, [4]Deepti Mousik, [2]Dr. S.R. Mansore, [3]Prof. R.B. Gaikwad

[1]Student, [2] professor, [3] professor, [4] professor

[1]Ujjain Engineering College, Ujjain,

[2]Ujjain Engineering College, Ujjain,

[3]Ujjain Engineering College, Ujjain,

[4]Ujjain Engineering College, Ujjain

Abstract

Image steganography has emerged as an important information security technique that enables confidential communication by concealing secret data within digital images. With the rapid growth of multimedia applications and cyber threats, researchers have developed a wide range of steganographic methods to improve embedding capacity, imperceptibility, robustness, and resistance against steganalysis. This review presents an overview of both traditional and modern image steganography techniques, including spatial-domain, transform-domain, adaptive, and deep learning-based approaches. The paper also discusses recent advancements in steganalysis, highlighting the role of artificial intelligence in detecting hidden information. A comparative analysis of representative techniques is provided to evaluate their strengths and limitations in terms of security, payload capacity, and visual quality. Finally, current research challenges and future directions are discussed to provide insights into the continued evolution of secure image steganography [20], [21].

Keywords: Image Steganography, Least Significant Bit (LSB), Adaptive Steganography, Deep Learning, Steganalysis, Information Security

1. Introduction

The rapid growth of digital communication and multimedia technologies has significantly increased the need for secure methods of information exchange. While cryptography protects the content of a message by converting it into an unreadable format, it does not conceal the existence of the communication itself. Image steganography addresses this limitation by embedding confidential information within digital images in a manner that remains visually imperceptible to human observers. As a result, steganography has become an important research area for applications requiring covert communication, digital rights protection, medical data security, military communication, and cloud-based information sharing [1], [4].

Over the past two decades, image steganography has evolved from simple spatial-domain techniques, such as Least Significant Bit (LSB) substitution, to more sophisticated transform-domain, adaptive, and artificial intelligence (AI)-based approaches. Traditional methods primarily focused on maximizing payload capacity while preserving image quality. However, modern techniques emphasize achieving a balance between embedding capacity, imperceptibility, robustness against image processing operations, and resistance to advanced steganalysis methods. Recent developments in deep learning, including convolutional neural networks, autoencoders, transformers, and invertible neural networks [7],[10], have further transformed the field by enabling intelligent and highly secure data hiding mechanisms.

Despite these advancements, several challenges remain. Increasing payload capacity without degrading visual quality, improving robustness against compression and image transformations, and defending against AI-driven steganalysis continue to be active research problems [20], [21]. Furthermore, the growing adoption of deep learning has introduced new opportunities as well as computational challenges, motivating researchers to explore more efficient and secure steganographic frameworks.

This review paper presents a comprehensive overview of modern image steganography techniques by examining the evolution of conventional spatial- and transform-domain methods alongside recent adaptive and deep learning-based approaches. In addition, recent developments in image steganalysis are discussed to highlight current detection capabilities and emerging security concerns. Finally, the paper summarizes existing challenges and identifies potential future research directions that may contribute to the development of more secure, robust, and intelligent image steganography systems [20],[22].

2. Traditional Image Steganography Techniques

Traditional image steganography techniques established the foundation for secure data hiding by embedding secret information into digital images while preserving their visual appearance. These methods are generally classified into spatial-domain and transform-domain approaches based on the location where the embedding process is performed. Spatial-domain techniques directly modify pixel values, offering high embedding capacity and low computational complexity. In contrast, transform-domain techniques embed information into frequency coefficients obtained through mathematical transformations, providing improved robustness against compression and common image processing operations. Although modern adaptive and deep learning-based methods have gained significant attention, traditional techniques remain widely used due to their simplicity, efficiency, and ease of implementation [3],[30].

2.1 Spatial Domain Techniques : Spatial-domain steganography embeds secret information by directly modifying the intensity values of image pixels. Among these methods, the Least Significant Bit (LSB) technique is the most widely adopted because of its simple implementation, low computational cost, and high embedding capacity. In LSB substitution, the least significant bits of selected pixels are replaced with the secret data, causing minimal visual distortion that is generally imperceptible to the human eye. However, conventional LSB methods are highly vulnerable to statistical steganalysis and image processing operations such as compression, filtering, and noise addition [4], [26], [30].

To overcome these limitations, several improved spatial-domain techniques have been proposed, including Pixel Value Differencing (PVD), edge-based embedding, adaptive LSB, and randomized pixel selection methods. These approaches intelligently select embedding regions according to local image characteristics, allowing higher payload capacity while reducing visual artifacts and improving resistance to detection. Despite these improvements, spatial-domain methods generally provide lower robustness than transform-domain approaches when subjected to significant image modifications [3], [30].

2.2 Transform Domain Techniques: Transform-domain steganography conceals secret information by embedding data into the frequency components of an image rather than directly modifying pixel values. This is typically achieved using mathematical transformations such as the Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Discrete Fourier Transform (DFT). By altering selected transform coefficients, these techniques offer greater robustness against image compression, filtering, and various signal processing operations while maintaining satisfactory visual quality [3], [15].

Compared with spatial-domain methods, transform-domain approaches generally provide enhanced security and resistance to steganalysis because the embedded information is distributed across the transformed image representation. Consequently, these techniques are widely employed in applications requiring reliable data hiding under image transmission and storage conditions. However, they usually involve higher computational complexity and lower embedding capacity, making implementation more challenging for real-time and resource-constrained environments. As a result, recent research has focused on combining transform-domain techniques with adaptive algorithms and artificial intelligence to further improve security, robustness, and embedding efficiency.

3. Modern Image Steganography Techniques

Recent advancements in artificial intelligence and deep learning have significantly transformed image steganography by enabling more intelligent, secure, and robust data hiding mechanisms. Unlike traditional approaches that rely on manually designed embedding rules, modern techniques employ adaptive algorithms and neural networks to automatically determine suitable embedding locations and optimize the balance between payload capacity, imperceptibility, and resistance to steganalysis. These methods have demonstrated remarkable improvements in visual quality and security, making them suitable for contemporary multimedia communication systems. Modern image steganography can be broadly categorized into adaptive steganography, deep learning-based steganography, and AI-based secure image hiding frameworks.

3.1 Adaptive Steganography : Adaptive steganography improves the security of data hiding by selecting embedding locations based on the characteristics of the cover image rather than using a fixed embedding strategy. These methods analyze image properties such as edges, textures, and noise levels to identify regions where modifications are less likely to be detected by the human visual system or statistical steganalysis techniques. Since textured and edge regions naturally contain greater intensity variations, embedding data in these areas minimizes visual distortion while maintaining higher payload capacity.

Several adaptive approaches combine image analysis with optimization techniques to determine the most appropriate embedding positions. By dynamically adjusting the embedding process according to local image content, adaptive steganography achieves better imperceptibility and enhanced resistance to steganalysis compared with conventional spatial-domain methods. As a result, adaptive embedding has become an important intermediate step between traditional steganography and modern AI-driven data hiding systems [17], [19].

3.2 Deep Learning-Based Steganography: Deep learning has revolutionized image steganography by enabling neural networks to learn optimal data hiding strategies directly from large image datasets. Instead of relying on manually designed embedding rules, deep learning models automatically extract image features and determine suitable embedding patterns that maximize visual quality while minimizing the probability of detection. Convolutional Neural Networks (CNNs), autoencoders, Generative Adversarial Networks (GANs), Vision Transformers (ViTs), and Invertible Neural Networks (INNs) have emerged as the most widely used architectures for secure image steganography [6]–[10].

Recent methods such as HiNet, DeepMIH, and invertible neural network-based frameworks simultaneously optimize image reconstruction and secret data recovery, achieving high embedding capacity with minimal visual distortion. These approaches demonstrate superior robustness against image processing operations and modern steganalysis methods while maintaining excellent reconstruction accuracy. Although deep learning-based techniques require higher computational resources and extensive training datasets, they have established a new benchmark for secure and intelligent image steganography [7]–[10].

3.3 AI-Based Secure Image Hiding Frameworks: Artificial intelligence has further enhanced image steganography by integrating machine learning, optimization algorithms, and intelligent decision-making into the embedding process. AI-based secure image hiding frameworks extend conventional deep learning approaches by incorporating adaptive optimization, feature selection, encryption, and hybrid embedding strategies to improve both security and robustness. These frameworks often combine multiple techniques, including neural networks, evolutionary algorithms, attention mechanisms, and image quality optimization, to achieve an effective balance between embedding capacity, imperceptibility, and resistance to steganalysis.

Recent research has introduced intelligent frameworks that employ invertible neural networks, multimodal learning, and hybrid optimization methods for secure image hiding. Such systems are capable of automatically learning robust embedding representations while preserving image quality and ensuring accurate recovery of the hidden information. As AI technologies continue to advance, secure image hiding frameworks are expected to become more adaptive, computationally efficient, and resilient against increasingly sophisticated steganalysis techniques, making them a promising direction for future research in image steganography [20]–[22].

4. Image Steganalysis

Image steganalysis is the process of detecting the presence of hidden information within digital images without prior knowledge of the embedding algorithm or secret key. As image steganography techniques have become increasingly sophisticated, steganalysis methods have also evolved to identify subtle modifications introduced during the embedding process. The primary objective of steganalysis is to distinguish between clean (cover) images and stego images by analyzing statistical properties, image features, or learned patterns. Modern steganalysis plays a crucial role in digital forensics, cybersecurity, law enforcement, and information security by preventing the misuse of steganographic techniques for illegal or malicious communication. Depending on the detection methodology, steganalysis techniques can be broadly classified into conventional methods, machine learning-based approaches, and deep learning-based techniques [5], [13].

4.1 Conventional Steganalysis: Conventional steganalysis techniques rely on statistical analysis and handcrafted image features to detect hidden information. These methods examine changes in pixel distributions, histogram characteristics, noise patterns, and higher-order statistical relationships that may result from the embedding process. Early techniques such as histogram analysis, chi-square attacks, RS analysis, and sample pair analysis proved effective against simple spatial-domain methods, particularly Least Significant Bit (LSB) substitution. Although these approaches are computationally efficient and easy to implement, their performance decreases significantly when applied to modern adaptive and deep learning-based steganographic algorithms, which introduce fewer detectable statistical artifacts [2], [5].

4.2 Machine Learning-Based Steganalysis: Machine learning-based steganalysis enhances detection accuracy by utilizing classifiers trained on discriminative image features extracted from both cover and stego images. Instead of relying solely on predefined statistical rules, these approaches employ feature extraction techniques followed by classifiers such as Support Vector Machines (SVM), Random Forests, Decision Trees, and Ensemble Learning models. Rich image models and high-dimensional feature representations have further improved the capability of machine learning algorithms to detect increasingly sophisticated steganographic methods. Although these techniques generally outperform conventional statistical approaches, their effectiveness largely depends on the quality of handcrafted features and the availability of representative training datasets [5], [13].

4.3 Deep Learning-Based Steganalysis: Deep learning-based steganalysis has emerged as the most effective approach for detecting hidden information in modern image steganography. Deep neural networks automatically learn hierarchical image features directly from training data, eliminating the need for manually designed feature extraction methods. Convolutional Neural Networks (CNNs), Residual Networks (ResNets), Vision

Transformers (ViTs), and hybrid deep learning architectures have demonstrated superior performance in identifying subtle embedding artifacts produced by advanced steganographic algorithms.

Recent deep learning-based steganalysis models achieve high detection accuracy even against adaptive and AI-driven steganography by learning complex spatial and frequency-domain representations. These models continue to improve through larger datasets, enhanced network architectures, and adversarial training strategies. However, they require substantial computational resources and large volumes of labeled training data, making their deployment more challenging than traditional detection techniques. Despite these limitations, deep learning has become the dominant direction for modern image steganalysis research due to its superior accuracy, scalability, and adaptability [20], [29].

5. Comparative Analysis

The performance of image steganography techniques is commonly evaluated using parameters such as embedding capacity, imperceptibility, robustness, computational complexity, and resistance to steganalysis. Traditional techniques are simple and computationally efficient but often provide limited security against modern detection methods. In contrast, adaptive, deep learning-based, and AI-driven approaches offer improved security and robustness by intelligently selecting embedding locations and learning optimal hiding strategies. Table 1 presents a comparative overview of the major categories of image steganography techniques [3], [20].

Technique	Embedding Capacity	Imperceptibility	Robustness	Computational Complexity	Resistance to Steganalysis
Spatial Domain	High	Moderate	Low	Low	Low
Transform Domain	Moderate	High	High	Moderate	Moderate
Adaptive Steganography	High	High	Moderate	Moderate	High
Deep Learning-Based	Very High	Very High	High	High	Very High
AI-Based Frameworks	Very High	Very High	Very High	High	Very High

Table 1. Comparison of Major Image Steganography Techniques

6. Challenges and Future Directions

Despite significant advancements in image steganography, several challenges remain that limit the practical deployment of modern data hiding techniques. One of the primary challenges is achieving an optimal balance between embedding capacity, imperceptibility, robustness, and security. Increasing the amount of hidden data often leads to greater image distortion and a higher risk of detection. Similarly, while deep learning-based methods provide superior performance, they require large training datasets, high computational resources, and considerable processing time, making them less suitable for resource-constrained environments.

Another important challenge is the rapid advancement of AI-driven steganalysis techniques, which are capable of detecting increasingly subtle embedding artifacts. As detection methods continue to improve, steganographic algorithms must become more adaptive and resilient to maintain secure communication. Furthermore, ensuring robustness against common image processing operations such as compression, resizing, cropping, and noise addition remains a critical research problem [5], [20].

Future research is expected to focus on developing lightweight and computationally efficient deep learning models that can be deployed on real-time and mobile applications. The integration of advanced artificial intelligence, attention mechanisms, transformer architectures, and hybrid optimization algorithms is likely to further enhance the security and robustness of image steganography. Additionally, research on explainable AI, multimodal data hiding, and quantum-resistant steganographic techniques may open new directions for secure information hiding in future digital communication systems [21], [22].

7. Conclusion

Image steganography has evolved from simple pixel-based data hiding techniques to advanced artificial intelligence-driven frameworks capable of providing high security, robustness, and embedding capacity. Traditional spatial- and transform-domain methods continue to offer efficient and practical solutions for many applications, while adaptive and deep learning-based approaches have significantly improved imperceptibility and resistance to modern steganalysis. At the same time, advances in steganalysis have driven the development of more secure and intelligent embedding techniques [7]–[10].

This review has presented an overview of traditional and modern image steganography methods, recent developments in steganalysis, and a comparative analysis of major approaches. The discussion also highlighted the current challenges and emerging research directions in this rapidly evolving field. As artificial intelligence continues to advance, future image steganography systems are expected to become more adaptive, computationally efficient, and resilient against sophisticated attacks, enabling secure and reliable information hiding for next-generation digital communication applications [20]–[22].

References

[1] N. Provos and P. Honeyman, "Hide and Seek: An Introduction to Steganography," IEEE Security & Privacy, vol. 1, no. 3, pp. 32–44, May–Jun. 2003.DOI: 10.1109/MSECP.2003.1203220

[2] J. Fridrich, M. Goljan, and R. Du, "Reliable Detection of LSB Steganography in Color and Grayscale Images," Proceedings of the ACM Workshop on Multimedia and Security, 2001, pp. 27–30.DOI: 10.1145/1232454.1232469

[3] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications. Cambridge University Press, 2009.DOI: 10.1017/CBO9780511576268

[4] N. F. Johnson and S. Jajodia, "Exploring Steganography: Seeing the Unseen," IEEE Computer, vol. 31, no. 2, pp. 26–34, 1998.DOI: 10.1109/2.650117

[5] J. Fridrich and J. Kodovsky, "Rich Models for Steganalysis of Digital Images," IEEE Transactions on Information Forensics and Security, vol. 7, no. 3, pp. 868–882, 2012.DOI: 10.1109/TIFS.2012.2190402

[6] J. Hayes and G. Danezis, "Generating Steganographic Images via Adversarial Training," Advances in Neural Information Processing Systems (NeurIPS), 2017.

[7] J. Jing, X. Deng, M. Xu, J. Wang, and Z. Guan, "HiNet: Deep Image Hiding by Invertible Network," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021.

[8] H. Lu, C. Luo, X. Zhang, and Z. Qin, "Large-Capacity Image Steganography Based on Invertible Neural Networks," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

[9] Z. Zhang, G. Fu, F. Di, and C. Qin, "Robust Invertible Image Steganography," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022.

[10] H. Bui, T. D. Nguyen, and H. Kim, "DeepMIH: Deep Invertible Network for Multiple Image Hiding," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 4, pp. 4972–4987, 2023.DOI: 10.1109/TPAMI.2022.3219556

[11] C. E. Shannon, "Communication Theory of Secrecy Systems," Bell System Technical Journal, vol. 28, no. 4, pp. 656–715, 1949.DOI: 10.1002/j.1538-7305.1949.tb00928.x

[12] A. Westfeld and A. Pfitzmann, "Attacks on Steganographic Systems," Proceedings of Information Hiding, 1999.DOI: 10.1007/3-540-48285-7_21

[13] J. Kodovsky, J. Fridrich, and V. Holub, "Ensemble Classifiers for Steganalysis of Digital Media," IEEE Transactions on Information Forensics and Security, vol. 7, no. 2, pp. 432–444, 2012.DOI: 10.1109/TIFS.2011.2175919

[14] G. J. Simmons, "The Prisoners' Problem and the Subliminal Channel," Advances in Cryptology (CRYPTO), 1983.DOI: 10.1007/978-1-4757-0602-4_5

[15] K. Solanki, N. Jacobsen, U. Madhow, B. S. Manjunath, and S. Chandrasekaran, "Robust Image-Adaptive Data Hiding Using Erasure and Error Correction," IEEE Transactions on Image Processing, vol. 13, no. 12, pp. 1627–1639, 2004.DOI: 10.1109/TIP.2004.837604

[16] V. Holub and J. Fridrich, "Designing Steganographic Distortion Using Directional Filters," Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS), 2012.DOI: 10.1109/WIFS.2012.6412655

[17] T. Pevný, T. Filler, and P. Bas, "Using High-Dimensional Image Models to Perform Highly Undetectable Steganography," Information Hiding, Lecture Notes in Computer Science, vol. 6387, Springer, 2010.DOI: 10.1007/978-3-642-16435-4_12

[18] T. Filler, J. Judas, and J. Fridrich, "Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes," IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 920–935, 2011.DOI: 10.1109/TIFS.2011.2134094

[19] V. Holub, J. Fridrich, and T. Denemark, "Universal Distortion Function for Steganography in an Arbitrary Domain," EURASIP Journal on Information Security, 2014.DOI: 10.1186/1687-417X-2014-1

[20] X. Zhang, S. Wang, Z. Qin, and C. Luo, "Deep Learning-Based Image Steganography: A Survey," IEEE Access, vol. 10, pp. 101234–101259, 2022.DOI: 10.1109/ACCESS.2022.3208097

[21] M. Ahmadi Bamanga, A. K. Babando, and M. A. Shehu, "Recent Advances in Steganography," IntechOpen, 2024.DOI: 10.5772/intechopen.1004521

[22] A. K. Op Dangi, S. Tandon, S. Deorari, and R. Kumar, "Steganography: Unveiling Techniques and Research Agenda," IntechOpen, 2024.DOI: 10.5772/intechopen.1005052

[23] R. Chandramouli, M. Kharrazi, and N. Memon, "Image Steganography and Steganalysis: Concepts and Practice," International Workshop on Digital Watermarking, 2003.DOI: 10.1007/978-3-540-39881-5_2

[24] C. Cachin, "An Information-Theoretic Model for Steganography," Information and Computation, vol. 192, no. 1, pp. 41–56, 2004.DOI: 10.1016/j.ic.2004.02.003

[25] R. J. Anderson and F. A. P. Petitcolas, "On the Limits of Steganography," IEEE Journal on Selected Areas in Communications, vol. 16, no. 4, pp. 474–481, 1998.DOI: 10.1109/49.668971

[26] J. Mielikäinen, "LSB Matching Revisited," IEEE Signal Processing Letters, vol. 13, no. 5, pp. 285–287, 2006.DOI: 10.1109/LSP.2006.870357

[27] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for Data Hiding," IBM Systems Journal, vol. 35, no. 3–4, pp. 313–336, 1996.DOI: 10.1147/sj.353.0313

[28] I. J. Cox, M. L. Miller, and J. A. Bloom, Digital Watermarking. Morgan Kaufmann, 2002.DOI: 10.1016/B978-155860714-6/50031-2

[29] Y. Wang and P. Moulin, "Optimized Feature Extraction for Learning-Based Image Steganalysis," IEEE Transactions on Information Forensics and Security, vol. 2, no. 1, pp. 31–45, 2007.DOI: 10.1109/TIFS.2006.890289

[30] M. Garg, J. S. Ubhi, and A. K. Aggarwal, "Steganography and its Current Advancements in Spatial Domain," EasyChair Preprint No. 2177, 2019.

[31] T. Filler and J. Fridrich, "Gibbs Construction in Steganography," IEEE Transactions on Information Forensics and Security, vol. 5, no. 4, pp. 705–720, 2010.DOI: 10.1109/TIFS.2010.2071879

[32] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically Undetectable JPEG Steganography: Dead Ends, Challenges, and Opportunities," Proceedings of the 9th ACM Multimedia & Security Workshop, 2007.DOI: 10.1145/1314534.1314572

[33] C. Qin, C. C. Chang, Y. Hsu, and L. T. Liao, "An Inpainting-Assisted Reversible Steganographic Scheme Using a Histogram Shifting Mechanism," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 7, pp. 1109–1118, 2013.DOI: 10.1109/TCSVT.2012.2234092