
Image Information Security Protection in the Big Data Era Based on Reversible Data Hiding in Encrypted Images

JunYi Lu

*School of Digital Economy, Nanning Vocational and Technical University, Nanning
530008, China*

E-mail: Lujunyi202512@163.com

Received 13 April 2026; Accepted 17 June 2026

Abstract

Massive image sharing and cloud-based storage in big data environments have significantly increased the risks of privacy leakage, unauthorized tampering, and insecure transmission. To address these issues, a reversible data hiding in encrypted images (RDHEI) framework integrating reserved room before encryption (RRBE) and vacating room after encryption (VRAE) is proposed for secure image information protection. The method combines high-bit plane prediction compression, key-controlled image encryption, and ciphertext-domain adaptive embedding to simultaneously achieve high-capacity data hiding, strict permission separation, and lossless image recovery. In addition, an error reverse reduction mechanism is introduced to suppress pixel distortion and improve recovery stability. Experimental evaluation is conducted on the UCID public image dataset containing 1338 uncompressed color images with rich texture and structural characteristics, including natural scenes, urban buildings, portraits, vehicles, plants, animals, and indoor environments. Experimental results on the UCID dataset demonstrate that the proposed method achieves an embedding rate of 0.99 bpp, a payload capacity of 524000 bits, Peak Signal-to-Noise Ratio (PSNR) of 52.14 dB,

Journal of Cyber Security and Mobility, Vol. 15.5, 1287–1310.

doi: 10.13052/jcsm2245-1439.1556

© 2026 River Publishers

and Structural Similarity Index Measure (SSIM) of 0.994. Meanwhile, the bit error rate is reduced to 0.028, the information entropy reaches 7.999, and the Number of Pixels Change Rate (NPCR) reaches 99.37%, indicating strong encryption randomness and resistance to differential attacks. Compared with Convolutional Neural Network-based Reversible Data Hiding (CNN-RDH) and Generative Adversarial Network-based Reversible Data Hiding (GAN-RDH) methods, the proposed framework achieves superior performance in embedding efficiency, recovery accuracy, visual quality, and computational efficiency. The proposed RDHEI framework provides an effective and scalable solution for secure image storage, privacy-preserving transmission, and encrypted multimedia sharing in large-scale big data applications.

Keywords: Reversible data hiding, encrypted images, information security protection, image encryption, big data.

1 Introduction

The popularization of mobile Internet and cloud computing environment has led to explosive growth of image data. Massive images are frequently collected, stored and transmitted in social media, telemedicine and other scenarios [1]. In an open network environment, images contain sensitive content such as identity information. Once leaked or tampered with, personal privacy and social security will be seriously affected. However, traditional image encryption technology relies on scrambling and diffusion mechanisms to enhance data confidentiality, but cannot flexibly embed authentication or traceability information in the encrypted state. The digital watermark method can achieve copyright protection and integrity verification, but it will cause irreversible distortion to the image, making it difficult to meet the application requirements for lossless recovery [2–4]. In response to these problems, Singh and Kumar focused on the risk of privacy leakage in network communication images, proposed a multi-level image encryption framework and improved authentication technology. This framework combined the public key encryption system and double chaos mapping to expand the key space and reduced the correlation between adjacent image pixels. It had good resistance to statistical analysis and brute force cracking capabilities [5]. Considering the insufficient research on three-dimensional medical image encryption, Liu and Xue suggested a three-dimensional encoding scheme that combined biometric keys and cubic box structures. The biometric key enhanced security attributes, and the cubic box introduced strong nonlinearity to make the

encryption process more complex. Tests showed that the scheme performed well in terms of key space scale and could effectively resist a variety of typical attacks [6]. Regarding parallel encryption of multiple images, Wang et al. put forward an encryption method grounded on computational generation of phase-only holograms and chaos models. This method used an improved Gerchberg-Saxton algorithm to obtain subsampled phase holograms, redistributed the phase information of multiple images in the form of spatial division, and used chaos transformation to generate ciphertext. The results showed that it could effectively prevent information leakage and improve encryption complexity and security [7]. Aiming at the security and efficiency issues in color image transmission and storage, Yao et al. raised a compression encryption method. It combined compressed sensing, Sudoku matrix structure and hyper-chaotic system. This method constructed a new hyper-chaotic map, used an improved neodymium optimization algorithm to adaptively adjust the compression threshold, and introduced Sudoku constraints and bidirectional diffusion strategies to improve the security performance of the algorithm. Results showed that this scheme performed well on multiple security indicators [8], but this method still has error propagation and low recovery stability. Although existing studies have improved image encryption security and reversible embedding performance to some extent, most methods still suffer from limited embedding flexibility, insufficient ciphertext-domain operability, error propagation during recovery, and high computational complexity. In addition, many traditional methods mainly focus on either encryption security or embedding performance, while lacking balanced optimization among embedding capacity, recovery accuracy, encryption randomness, and computational efficiency. These limitations reduce their adaptability in large-scale and complex big data image environments. Therefore, the main problem addressed in this study is how to achieve a balanced optimization among embedding capacity, reversible recovery accuracy, encryption security, and computational efficiency in encrypted image environments. Existing methods often improve one aspect while sacrificing others, which limits their applicability in large-scale big data scenarios. To solve this problem, a reversible data hiding in encrypted images (RDHEI) framework integrating reserved room before encryption (RRBE) and vacating room after encryption (VRAE) strategies is proposed to simultaneously improve embedding efficiency, recovery stability, and ciphertext-domain security. Therefore, an image information security protection model in the big data era grounded on RDHEI was proposed. This model constructed redundant space through high-bit prediction compression and performed adaptive embedding through

the ciphertext domain reconstruction mechanism. The research aimed to improve the randomness of image encryption, reduce recovery errors and optimize computing efficiency, so as to provide scalable solutions for large-scale image security applications. The key novelty of this research lies in the construction of a dual-stage reversible data hiding (RDH) framework that jointly integrates RRBE and VRAE. Unlike conventional RDH or standalone RDHEI methods, the proposed model simultaneously achieves high-capacity embedding, strict key-separated access control, and lossless image recovery within the encrypted domain. In addition, the study introduces high-bit plane prediction compression and ciphertext-domain adaptive embedding strategies to improve embedding efficiency while reducing pixel distortion and error propagation. The proposed mechanism further enhances encryption randomness, recovery reliability, and resistance to differential attacks, thereby providing a scalable and secure solution for large-scale image privacy protection in big data environments. Compared with existing RDH and RDHEI methods reported in the literature, the proposed approach exhibits several advantages. First, conventional Convolutional Neural Network-based Reversible Data Hiding (CNN-RDH) and Generative Adversarial Network-based Reversible Data Hiding (GAN-RDH) methods mainly rely on deep feature reconstruction or adversarial generation, which often introduce reconstruction instability and error propagation during reversible recovery. In contrast, the proposed model adopts deterministic prediction compression and ciphertext-domain adaptive embedding, thereby achieving lower bit error rates and higher recovery accuracy. Second, many traditional RDHEI frameworks only employ either RRBE or VRAE independently, resulting in limited embedding capacity or insufficient ciphertext flexibility. The proposed hybrid RRBE-VRAE structure simultaneously preserves embedding redundancy before encryption and releases additional embedding space after encryption, significantly improving payload capacity and embedding efficiency. Third, the proposed method achieves better encryption randomness and resistance to differential attacks, as demonstrated by higher information entropy and NPCR values. Furthermore, the model maintains lower computational complexity and shorter processing time while preserving high visual quality and structural similarity. These advantages indicate that the proposed framework provides a more balanced solution among security strength, embedding capability, reversibility, and computational efficiency.

The importance of the current work is reflected in both theoretical and practical aspects. (1) From a theoretical perspective, the proposed RDHEI framework extends conventional RDH research from plaintext-domain

embedding to a hybrid encrypted-domain collaborative protection mechanism, thereby improving the coordination between embedding capacity, recovery accuracy, and encryption security. The integration of RRBE and VRAE further provides a new strategy for balancing reversible recovery and ciphertext-domain operability. (2) From a practical perspective, the proposed method is suitable for large-scale image storage, cloud transmission, medical image management, intelligent surveillance, and cross-platform privacy sharing in big data environments. Since the model supports strict permission separation and complete lossless recovery, it can effectively reduce privacy leakage risks while maintaining data integrity and traceability. Therefore, the proposed framework has potential application value in secure multimedia communication and privacy-preserving data management systems.

2 Methods

2.1 Image Information Security Protection Model Based on RDH

In the big data environment, image data is large, diverse, and spread frequently. Only encryption or access control methods can no longer simultaneously meet the requirements for data integrity, traceability, and recoverability. RDH can embed relevant information without distortion to achieve image information security protection [9, 10]. This method mines the spatial correlation and statistical redundancy between image pixels, and can still accurately restore the original image after embedding certification marks, copyright watermarks or index information, avoiding the permanent distortion problem caused by the use of traditional watermark technology [11]. In scenarios requiring high image content integrity, such as medical imaging and high-precision remote sensing, constructing an image information security protection model with RDH enables the embedding and identity authentication of sensitive information. Meanwhile, as illustrated in Figure 1, this model also ensures controllability and reliability throughout data management and storage.

From Figure 1, the RDH process includes preprocessing, predictive modeling, error expansion embedding, information extraction, and lossless recovery. The original carrier image is denoted as I , the pixel position as (i, j) , and the corresponding pixel value as $I(i, j)$. In the preprocessing module, local correlation is enhanced through sorting or smoothing operations to obtain the processed image I_p . Then, a prediction model based on neighborhood $N_{i,j}$ is constructed to obtain the pixel prediction value, whose

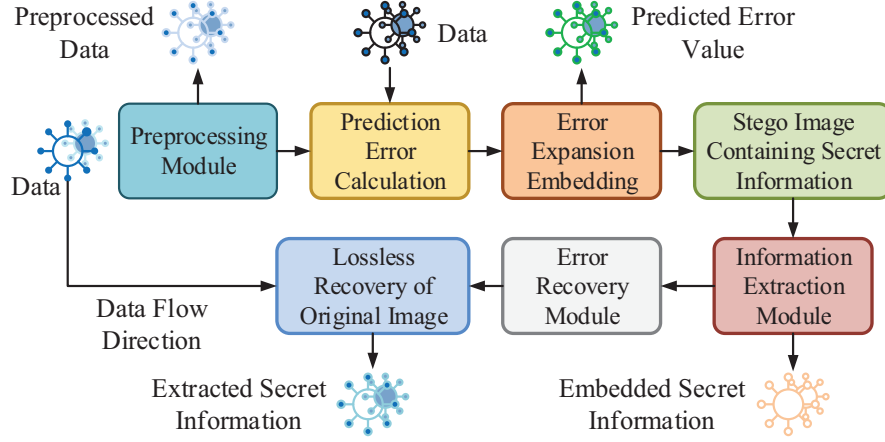


Figure 1 RDH flow chart.

expression is shown in Equation (1).

$$\hat{I}(i, j) = f(\mathcal{N}_{i,j}) \quad (1)$$

where $\hat{I}(i, j)$ is the predicted pixel value, $\mathcal{N}_{i,j}$ represents the neighborhood set of pixel (i, j) , and $f(\cdot)$ is the linear or weighted prediction function. The prediction error is defined as shown in Equation (2).

$$e(i, j) = I(i, j) - \hat{I}(i, j) \quad (2)$$

where $e(i, j)$ is the prediction residual, which is used to describe the difference between the real pixels and the predicted values. In the error expansion embedding stage, the bits $b \in \{0, 1\}$ to be embedded are integrated into the prediction error to form an expanded error, whose expression is shown in Equation (3).

$$e'(i, j) = 2e(i, j) + b \quad (3)$$

where $e'(i, j)$ is the expanded error value, and b is the embedded secret bit. The dense pixels are reconstructed by Equation (4).

$$I'(i, j) = \hat{I}(i, j) + e'(i, j) \quad (4)$$

where $I'(i, j)$ represents the dense image pixel. In the information extraction stage, the embedded bits are restored through modular operation, and its expression is shown in Equation (5).

$$b = e'(i, j) \bmod 2 \quad (5)$$

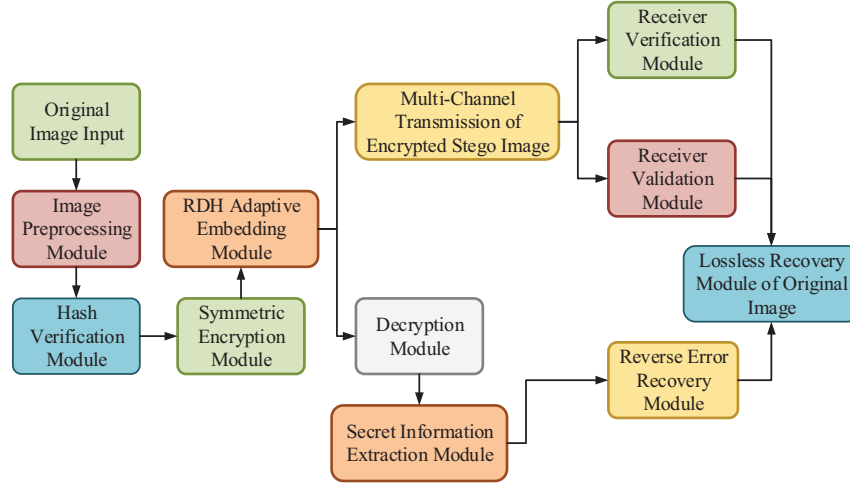


Figure 2 Overall framework of image information security protection based on RDH.

The original error is restored from the prediction residual, and its expression is shown in Equation (6).

$$e(i, j) = \left\lfloor \frac{e'(i, j)}{2} \right\rfloor \quad (6)$$

Through the original error, the original pixel value is reconstructed [12]. An image encryption model is established through RDH, and its structure is shown in Figure 2

From Figure 2, the process includes data input, preprocessing, integrity verification, encryption embedding, multi-channel transmission, receiver processing, and lossless recovery. First, the original image input system enters the image preprocessing module to perform format standardization, feature improvement or redundant space analysis on the image; image summary information is generated through hash verification to make sure the data's integrity and verifiability. Meanwhile, it is encrypted via the symmetric encryption module to achieve pixel-level privacy protection and ensure secure transmission over open networks. The encrypted image then enters the RDH adaptive embedding module, which embeds authentication marks or additional information in the ciphertext domain to form a confidential encrypted image and support multi-channel transmission. The receiving end performs verification and verification to determine whether the data is complete and consistent, and then decrypts and extracts secret information in sequence to

restore the embedded data; the original image is losslessly restored through error reverse restoration.

2.2 Image Information Security Protection Model Based on RDHEI

In the image information security protection model based on RDH, information embedding and lossless recovery are completed in the plaintext domain. Although the integrity and reversibility of the image content can be guaranteed, there is still a risk of privacy leakage in an open network environment or cloud storage scenario. To further improve the security level of the system, it is necessary to integrate the data hiding mechanism and image encryption technology to build a security protection framework with both confidentiality and recoverability [13, 14]. As a result, the RDHEI model introduces an encryption process grounded on the traditional RDH structure, so that the image is in a ciphertext state before embedding, effectively isolating the image content from the external access environment [15]. Meanwhile, through reserved space or adaptive embedding strategy, additional information can be embedded without destroying the randomness of the encryption structure, and secret data extraction and accurate recovery of the original image can be achieved at the receiving end. This study employs a modeling approach that combines the RDHEI algorithm with the Reserving Room Before Encrypted (RRBE) variant and the RDHEI algorithm with the Vacating Room After Encrypted (VRAE) variant. In the RDHEI framework, high-bit plane prediction compression refers to predicting and compressing the high-order effective information of image pixels to generate reversible embedding redundancy while preserving important image structures. The ciphertext-domain reconstruction mechanism means that secret information is embedded and extracted directly in the encrypted image domain without exposing the original image content, thereby improving privacy protection during transmission and storage. Meanwhile, the two-stage space reservation strategy combines RRBE and VRAE to jointly improve embedding flexibility and recovery reliability. This collaborative mechanism enables the encrypted image to maintain both high embedding capacity and stable reversible reconstruction performance. The RDHEI algorithm process based on RRBE is shown in Figure 3.

As shown in Figure 3, on the content owner side, the original image first enters the free-space module. Through operations such as pixel rearrangement, redundant compression, or embeddable area reservation, sufficient

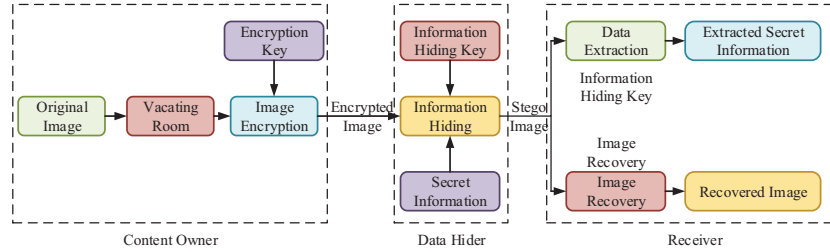


Figure 3 RDHEI algorithm flow based on RRBE.

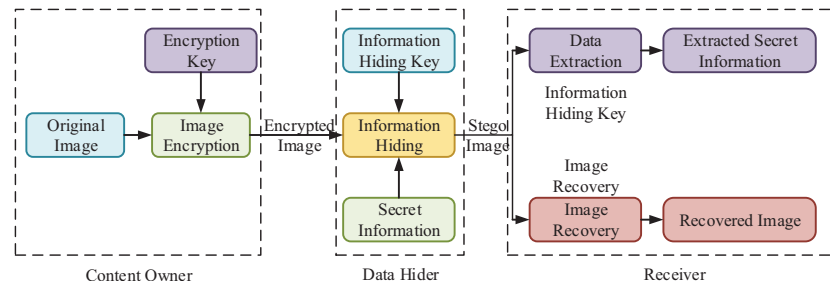


Figure 4 VRAE-based RDHEI algorithm flow.

available capacity is created for subsequent data hiding. Subsequently, the preprocessed image is encrypted using an encryption key to generate an encrypted image, which is then transmitted to the data embedder. Secondly, on the data embedder side, without knowing the content of the original image, the embedder embeds secret information into the encrypted image with reserved space with an independent information hiding key to form a secret image, thereby achieving data hiding and content isolation [16, 17]. Finally, on the receiver side, data extraction and image recovery operations can be performed separately depending on the type of key held. If you only have the information hiding key, you can extract secret information without knowing the image content. If you only have the encryption key, you can decrypt and restore the image but cannot extract the hidden data. If you have both types of keys, you can achieve secret information extraction and lossless recovery of the original image. The single RDHEI algorithm based on RRBE has insufficient performance, so this algorithm is combined with the RHEI algorithm based on VRAE. The process of the VRAE-based RHEI algorithm is shown in Figure 4.

From Figure 4, first, on the content owner side, the original image is directly encrypted using the encryption key without structural preprocessing,

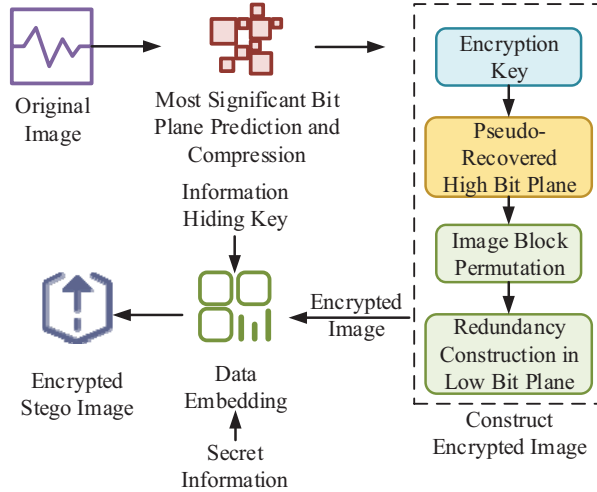


Figure 5 Image information security protection model based on RRBE-VRAE.

and the encrypted image is generated and transmitted to the data embedder. Subsequently, on the data embedder side, without accessing the original image content, the embedder analyzes the encrypted image using an independent information hiding key. By reorganizing bit planes in the ciphertext domain or adjusting part of the pixel space structure, it releases embeddable capacity and embeds secret information to form a secret image. Since the space-making operation occurs after encryption, this method does not need to reserve redundancy in the plaintext stage, but it has higher requirements on the operability of the ciphertext structure [18, 19]. Finally, on the receiver side, data extraction and image recovery operations can be performed depending on the key held. If you only have the information hiding key, you can extract the secret information but cannot restore the original image. if you only have the encryption key, you can decrypt the encrypted image and restore a visually acceptable image. When you have both types of keys, you can achieve accurate extraction of secret information and complete lossless recovery of the original image. The model after merging the two algorithms is shown in Figure 5.

From Figure 5, first, the original image I undergoes high-bit plane prediction and compression processing, and the pixel $I(i, j)$ is decomposed into a high-bit plane $H(i, j)$ and a low-bit plane $L(i, j)$, and its expression is shown in Equation (7).

$$I(i, j) = H(i, j) \cdot 2^k + L(i, j) \quad (7)$$

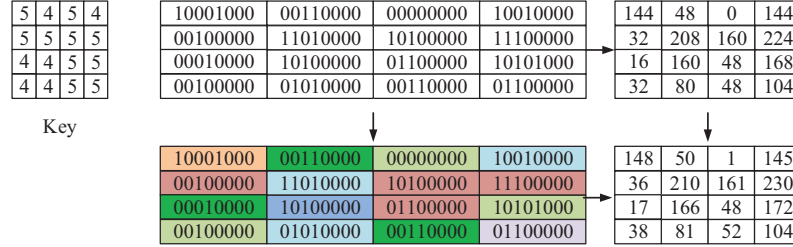


Figure 6 Schematic diagram of constructing redundancy to generate encrypted image.

where k is the number of low-order plane bits, $H(i, j)$ represents the high-order effective information, and $L(i, j)$ is the low-order redundant part. Then, prediction modeling is performed on the high-level plane to obtain the prediction value $\hat{H}(i, j)$, and its prediction error is shown in Equation (8) [20].

$$e(i, j) = H(i, j) - \hat{H}(i, j) \quad (8)$$

where $e(i, j)$ represents the high-bit plane prediction residual, and space reservation is achieved by statistically compressing it. Entering the encryption stage, under the control of the encryption key K_e , the image block is scrambled, and its expression is shown in Equation (9).

$$I_e = \mathcal{P}_{K_e}(I) \quad (9)$$

where $\mathcal{P}_{K_e}(I)$ denotes the key-based pixel scrambling function and I_e denotes the encrypted image, in which low-order structural redundancy is preserved for subsequent embedding. The schematic diagram of its structural redundancy to generate encrypted images is shown in Figure 6.

From Figure 6, the original pixel blocks and the corresponding key matrix are first given, and the pixel blocks are grouped and mapped under the control of the key. Meanwhile, the random matrix and the permutation matrix are combined to rearrange the pixel bit structure. The pixel values are then converted into binary form, and the low-order bits are controllably adjusted while the high-order effective information remains largely unchanged. This process generates exploitable structural redundancy with minimal impact on the visual quality of the image. The modified bits are marked in different colors in the Figure 6. These positions constitute the reserved space for subsequent data embedding. After matrix operations and key modulation, the updated binary sequence is recombined into new pixel values to form an encrypted image block. The pixel matrix after structural redundancy processing changes regularly in value, but the overall grayscale

distribution is still within a reasonable range, ensuring that the statistical characteristics of the image are not seriously damaged. Through bit-level reconstruction and key-driven structural transformation, this mechanism generates an embedding space simultaneously during the encryption process. It not only maintains the randomness and security strength of the encrypted image, but also provides support for subsequent secret information hiding and lossless recovery. Finally, under the control of the information hiding key K_h , the secret bit b is embedded in the reserved space to form a secret-carrying encrypted image, whose expression is shown in Equation (10).

$$I_{em} = I_e \oplus \Phi_{K_h}(b) \quad (10)$$

where $\Phi_{K_h}(b)$ represents the embedding mapping function generated by the hidden key, and $\oplus$ is the bitwise embedding operation. This model achieves high-capacity embedding and strict permission separation through a joint mechanism of reserved space before encryption and embedded data after encryption.

3 Results

3.1 Performance of Image Information Security Protection Model based on RDHEI

This study used a computer equipped with Windows 11 operating system and 32 GB of memory, equipped with Intel Core i5-6300 CPU. The data set selected the UCID public data set, which contained 1338 uncompressed color images, all stored in TIFF lossless format, with a resolution of 512×384 pixels and a color space of 24-bit RGB. The image content came from real shooting scenes, covering natural scenery, urban architecture, indoor scenes, people, vehicles, plants, animals and daily life objects, with rich texture structure and color distribution characteristics. Some images in the data set had similar scenes or similar objects in content, but different shooting angles, lighting conditions, and composition methods, which facilitated testing of the algorithm's distinguishing ability and robustness under similar content. Since the image was not compressed, its pixel values remained in their original acquired state, which made it suitable for pixel-level analysis tasks such as RDH, image encryption, embedding capacity evaluation, and lossless recovery performance testing. During the experiment, all models were trained and tested under the same hardware and dataset conditions to ensure comparison fairness. The maximum iteration number was set to 80, and the

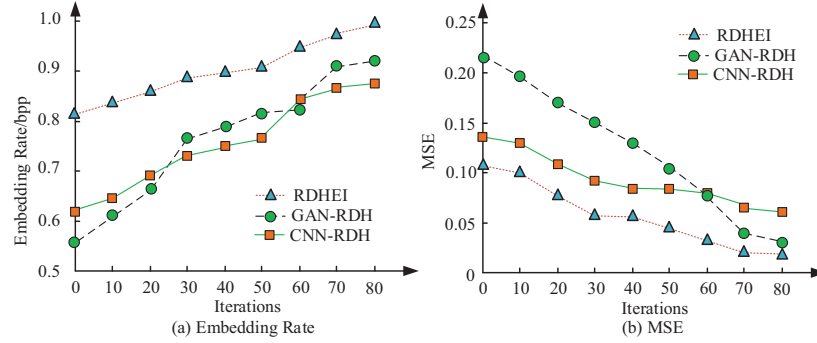


Figure 7 Analysis of (a) embedding rate and (b) mean square error of each model.

iteration process was terminated early when the change in loss value between two consecutive iterations was less than 10^{-5} . For the proposed RDHEI framework, the high-bit plane compression threshold was set to 4 bits, and the adaptive embedding strength parameter was initialized to 0.35 according to preliminary validation experiments. For CNN-RDH, the convolution kernel size was set to 3×3 , the learning rate was initialized to 0.001, and the batch size was set to 32. For GAN-RDH, both generator and discriminator networks were trained using the Adam optimizer with an initial learning rate of 0.0002 and momentum coefficients of 0.5 and 0.999. All comparison models were tuned using the same training dataset and evaluation protocol to ensure consistent optimization conditions. Each experiment was repeated five times, and the average results were reported to reduce randomness and improve statistical reliability. Comparative models included CNN-RDH and GAN-RDH.

Figure 7(a) shows the changing trend of the embedding rate of different models as the number of iterations increases. As the iterations increased from 10 to 80, the embedding rates of the three methods all showed a steady upward trend. Among them, RDHEI always maintained the highest level. The embedding rate was about 0.83 bpp at 10 iterations and close to 1.0 bpp at 80 iterations. The overall improvement was obvious. GAN-RDH increased from approximately 0.61 bpp to 0.91 bpp, while CNN-RDH increased from approximately 0.64 bpp to 0.87 bpp. Figure 7(b) shows the trend of the mean square error of the model changing with the number of iterations. The mean square error (MSE) of RDHEI was about 0.10 at 10 iterations, and dropped to about 0.02 at 80 iterations, which was the largest decrease; GAN-RDH dropped from about 0.20 to 0.03, and CNN-RDH dropped from 0.13 to about 0.06. The reason why RDHEI had lower error was that its

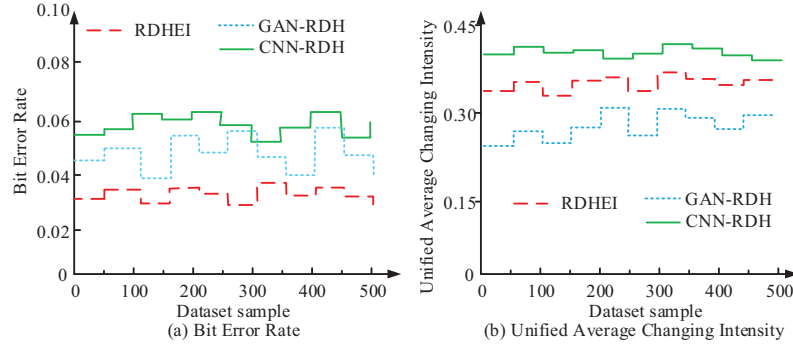


Figure 8 Analysis of (a) bit error rate and (b) unified average changing intensity of each model.

embedding process was based on reversible error modeling and redundant construction, which can accurately control the range of pixel modification. However, the generative method had pixel fluctuations and reconstruction errors in adversarial training, resulting in higher early distortion. In a word, RDHEI achieved a better balance between capacity and distortion control. The performance of each model was further analyzed. The results are shown in Figure 8.

Figure 8(a) demonstrates the changes in bit error rates of different models on multiple data samples. The overall bit error rate of RDHEI was the lowest, basically stable at around 0.03, and even lower than 0.03 at some samples, with a small fluctuation range. The bit error rate of GAN-RDH was mostly between 0.04 and 0.05, while that of CNN-RDH was generally between 0.055 and 0.065, with individual samples approaching 0.06 or more. The reason for the low bit error rate of RDHEI was that its embedding and recovery process was based on reversible structure modeling and key separation mechanism, and the data extraction process relied on deterministic mapping, so the error propagation was small. This was mainly because the proposed framework employed reversible prediction-error modeling and deterministic ciphertext-domain embedding mechanisms. The reserved redundancy generated through high-bit plane prediction compression reduced unnecessary pixel modification during the embedding process, thereby effectively suppressing error accumulation and propagation. In addition, the RRBE-VRAE collaborative structure enabled stable embedding space allocation before and after encryption, which improved the consistency of hidden data extraction and image reconstruction under different image textures and grayscale distributions. The independent key-separation mechanism further reduced

interference between encryption and embedding operations, thereby enhancing recovery reliability. By contrast, GAN-RDH relied on adversarial reconstruction and iterative generation, which could introduce local instability and reconstruction deviations during the recovery stage, while CNN-RDH was more sensitive to complex texture variations and local prediction errors. Therefore, RDHEI maintained lower bit error rate values and better stability across different test samples. Figure 8(b) shows the changing trend of the uniform average change intensity. GAN-RDH had the lowest overall value, roughly distributed between 0.27 and 0.31, RDHEI was at a medium level, mostly between 0.32 and 0.35, and CNN-RDH had the highest value, approximately between 0.38 and 0.43. A higher change intensity meant a greater overall perturbation of the pixel. RDHEI maintained a high change intensity while still having a low bit error rate, indicating that its pixel change distribution was more uniform and controlled, which was beneficial to improving encryption randomness and anti-difference capabilities. RDHEI achieved a more reasonable balance between security and reliability. Each model's comprehensive performance was analyzed. The results can be seen in Table 1.

From Table 1, RDHEI demonstrated superior comprehensive performance in embedding capability, image quality, recovery reliability, security strength, and computational efficiency. In terms of embedding performance, the embedding rate of RDHEI reached 0.98 bpp, which was approximately 7.7% higher than GAN-RDH and 12.6% higher than CNN-RDH. Meanwhile, the payload capacity increased to 512000 bits, indicating that the proposed RRBE-VRAE collaborative framework could provide more effective embedding redundancy and higher space utilization efficiency. In terms of image

Table 1 Comparison results of comprehensive performance of different models

Model	RDHEI	GAN-RDH	CNN-RDH	Average
Embedding Rate/bpp $\uparrow$	0.98	0.91	0.87	0.92
Payload Capacity/bits $\uparrow$	512000	468000	442000	474000
PSNR/dB $\uparrow$	51.62	47.35	44.92	47.96
SSIM $\uparrow$	0.992	0.981	0.973	0.982
MSE $\downarrow$	0.021	0.038	0.061	0.040
BER $\downarrow$	0.031	0.047	0.058	0.045
Information Entropy $\uparrow$	7.998	7.985	7.972	7.985
NPCR/% $\uparrow$	99.24	98.73	98.11	98.69
Unified Average Changing Intensity (UACI) $\uparrow$	0.334	0.298	0.412	0.348
Runtime/s $\downarrow$	0.82	1.36	1.12	1.10

quality preservation, the PSNR of RDHEI reached 51.62 dB, significantly higher than the comparison models, while the SSIM reached 0.992, demonstrating that the proposed reversible embedding mechanism could effectively preserve structural details and visual consistency even under high-capacity embedding conditions. Meanwhile, the MSE was only 0.021, which reflected that the proposed high-bit plane prediction compression strategy reduced unnecessary pixel modification during the embedding process. Regarding recovery reliability, the Bit Error Rate (BER) of RDHEI was only 0.031, which was substantially lower than GAN-RDH and CNN-RDH, indicating that the proposed deterministic embedding and reverse recovery mechanisms effectively suppressed error propagation during ciphertext extraction and image reconstruction. For security evaluation, the information entropy of RDHEI reached 7.998, which was very close to the theoretical ideal value of 8, demonstrating high encryption randomness. Meanwhile, the NPCR reached 99.24%, indicating strong resistance against differential attacks. Although the UACI of CNN-RDH was relatively high, its recovery stability and image quality were significantly lower, suggesting that excessive pixel perturbation may negatively affect reversible reconstruction performance. In addition, RDHEI required only 0.82 s for processing, which was faster than GAN-RDH and CNN-RDH. This result demonstrated that the proposed framework not only improved security and embedding capacity, but also maintained favorable computational efficiency for large-scale image security applications.

3.2 Simulation Performance Analysis

To further validate the model's performance, simulation analysis was used to further test the performance. The results are shown in Figure 9.

Figure 9(a) is a comparison chart of the structural similarity index of various models on six groups of test samples. DHRI is basically the best, with most test values above 0.97, and the sixth group of samples is above 1.0, which has strong structural retention. GAN-RDH SSIM is between 0.85 and 0.95, and the fifth group dropped to 0.84, indicating that there is a certain deviation in structural reconstruction under complex texture conditions. The overall SSIM distribution of CNN-RDH ranged from 0.80 to 0.92, and the fourth sample decreased to approximately 0.78, indicating relatively weak structural stability. This is mainly because CNN-based reversible embedding methods are sensitive to local prediction errors and spatial texture variations. When processing images with complex geometric structures or highly

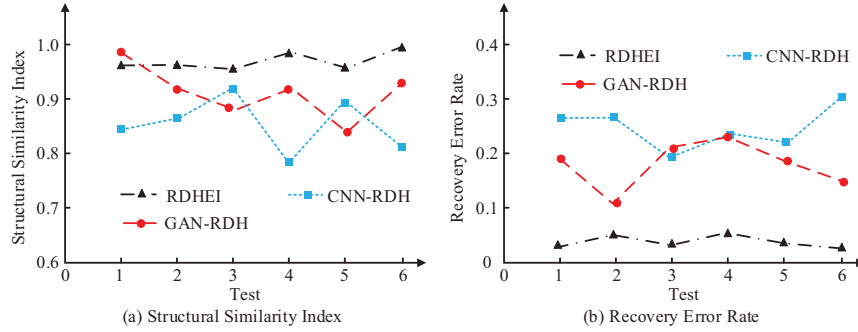


Figure 9 Analysis of (a) structural similarity index and (b) recovery error rate of each model.

correlated texture patterns, convolutional feature extraction may fail to preserve fine-grained edge continuity, thereby reducing structural similarity after recovery. Figure 9(b) is a comparison chart of recovery error rates. The error rates recovered by RDHEI in all tests are below 0.05, with the lowest value close to 0.02, and the recovery accuracy is the highest. The GAN-RDH error rate is between 0.12 and 0.23, the CNN-RDH error rate is between 0.20 and 0.30, and it reaches more than 0.30 on the sixth group. The error is very large. The higher recovery error rate indicated that there is error propagation in the model during the embedding or decoding process. Overall, RDHEI is more robust in terms of structure preservation and recovery accuracy. This study selected six representative image data types for testing to comprehensively test the model's adaptability under different content characteristics. Data type A was a natural landscape image, including large smooth areas and gradient structures such as mountains, rivers, and skies, and the texture was relatively uniform. Data type B was an urban architectural image, which contained a large number of regular edges and geometric structures, and had obvious linear texture characteristics. Data type C was a portrait image, which contained facial details and skin color gradient areas, and had high requirements for structure maintenance and visual quality. Data type D was a vehicle-type image with complex metal reflection and high-contrast edge features. Data type E was plant and animal images with rich textures, dense local details, and strong spatial correlation. Data type F was an indoor scene image, including furniture, walls and multi-scale texture structures, with complex grayscale distribution. The results are shown in Figure 10.

Figure 10(a) shows the information entropy performance of different models under five data types. RDHEI achieved the highest entropy value in all data types, which was close to 7.98 ± 0.03 in data Type E, which is

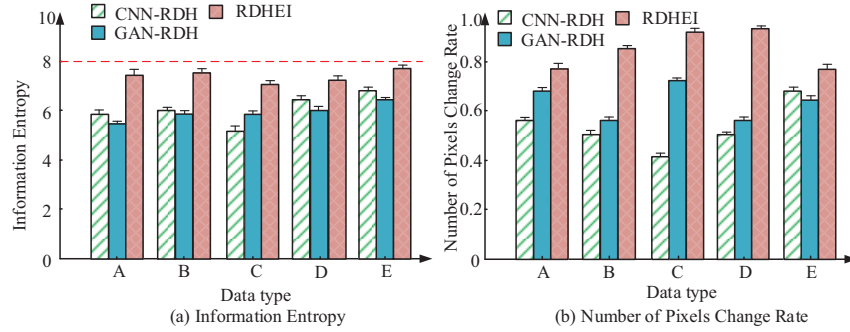


Figure 10 Analysis of each model's (a) information entropy and (b) pixel change rate.

significantly better than CNN-RDH of about 6.70 ± 0.11 and GAN-RDH of about 6.90 ± 0.08 . Overall, the entropy values of RDHEI in the five types of data from A to E are all stable above 7.50 ± 0.05 , while the other two methods are mostly distributed between 5.50 ± 0.14 and 7.00 ± 0.10 . The closer the information entropy is to 8, the closer the image grayscale distribution is to the ideal random state, and the stronger the resistance to statistical analysis. Figure 10(b) shows the comparison result of pixel change rate. RDHEI maintained a high level in all types of data, for example, close to 0.90 ± 0.02 in data Types C and D, which is significantly higher than CNN-RDH in Type C, about 0.40 ± 0.05 , and GAN-RDH, about 0.70 ± 0.03 . A higher pixel change rate means greater sensitivity to the original image, which is beneficial to improving the ability to resist differential attacks. RDHEI has more obvious advantages in randomness and attack resistance. Simulation analysis was used for testing. The results are shown in Table 2.

From Table 2, the proposed RDHEI framework maintained stable and superior performance under simulation testing conditions. The embedding capacity reached 0.99 bpp and the total payload increased to 524000 bits, indicating that the collaborative reserved-space strategy effectively improved ciphertext-domain embedding efficiency and data carrying capability. In terms of visual quality and structural preservation, RDHEI achieved a PSNR of 52.14 dB and an SSIM of 0.994, both of which are significantly higher than those of GAN-RDH and CNN-RDH. These results demonstrate that the proposed embedding mechanism can preserve image texture information and structural consistency while maintaining high embedding capacity. Furthermore, MSE is reduced to 0.018, reflecting more accurate pixel-level recovery performance. For reversible recovery evaluation, the bit error rate of RDHEI is only 0.028 and recovery accuracy reached 100%, indicating that

Table 2 Simulation performance test

Performance Metric	RDHEI	GAN-RDH	CNN-RDH
Embedding Capacity/bpp	0.99	0.92	0.88
Total Payload bits	524000	472000	451000
Peak Signal-to-Noise Ratio/dB	52.14	48.02	45.37
Structural Similarity Index	0.994	0.983	0.975
Mean Squared Error	0.018	0.034	0.057
Bit Error Rate	0.028	0.044	0.056
Information Entropy	7.999	7.988	7.976
Pixel Change Rate/%	99.37	98.84	98.29
Unified Average Changing Intensity	0.336	0.301	0.409
Recovery Accuracy/%	100.0	98.6	97.8
Processing Time/s	0.79	1.42	1.18

the proposed reverse error reduction strategy effectively minimized reconstruction distortion and ensured complete extraction of hidden information. By contrast, GAN-RDH and CNN-RDH exhibit relatively higher bit error rate and lower recovery accuracy due to instability in feature reconstruction and local prediction errors. In terms of encryption security, RDHEI achieves an information entropy of 7.999 and a pixel change rate of 99.37%, indicating that the encrypted image has strong randomness and high sensitivity to plaintext changes. These characteristics enhance the model's resistance to statistical analysis and differential attacks. Additionally, the processing time of RDHEI is only 0.79 s, which is significantly shorter than that of GAN-RDH. This demonstrates that the proposed framework achieves a more balanced optimization among embedding efficiency, image quality, reversible recovery capability, encryption security, and computational complexity.

4 Conclusion

To address the problems of privacy leakage, insecure transmission, and reversible recovery failure of image data in big data environments, an RDHEI framework integrating RRBE and VRAE was proposed. The framework combined high-bit plane prediction compression, key-controlled image scrambling, ciphertext-domain adaptive embedding, and reverse error recovery strategies to achieve high-capacity embedding, strict permission separation, secure transmission, and lossless image reconstruction within the encrypted domain. Experimental and simulation results demonstrated that the proposed RDHEI framework achieved superior overall performance compared with

CNN-RDH and GAN-RDH methods. The embedding rate reached 0.99 bpp, while the payload capacity increased to 524000 bits. Meanwhile, the PSNR reached 52.14 dB and the SSIM reached 0.994, indicating that the encrypted embedding process maintained high visual quality and strong structural consistency. The mean squared error was reduced to 0.018 and the bit error rate decreased to 0.028, reflecting stable reversible recovery capability and low error propagation. In terms of security performance, the information entropy reached 7.999 and the NPCR reached 99.37%, demonstrating strong encryption randomness and resistance to differential attacks. In addition, the overall processing time was only 0.79 s, which verified the computational efficiency and practical applicability of the proposed framework. Comprehensive analysis indicates that the proposed method achieves a more balanced optimization among embedding capacity, image quality, recovery reliability, encryption security, and computational efficiency. The framework therefore provides an effective solution for secure image storage, privacy-preserving cloud transmission, encrypted multimedia sharing, and sensitive image management in large-scale big data applications. Nevertheless, the current study mainly focuses on static image scenarios, and the adaptability of the framework under ultra-large-scale datasets and real-time multimedia transmission environments still requires further investigation. Future work will combine lightweight optimization, parallel computing strategies, and multimodal security protection mechanisms to improve real-time processing capability and expand the framework to video, medical imaging, and cross-platform intelligent data security applications.

Although the proposed RDHEI framework achieved favorable performance on the UCID dataset, the current experiments were mainly conducted under standard image conditions. Different image resolutions, compressed transmission environments, and real-world noisy scenarios may further influence embedding stability and reversible recovery performance. In addition, the present study mainly focused on encryption quality and reversible embedding evaluation, while robustness analysis under practical attack scenarios such as Gaussian noise interference, JPEG compression, cropping attacks, and partial data loss has not yet been comprehensively investigated. Nevertheless, the proposed ciphertext-domain adaptive embedding mechanism and reverse error reduction strategy demonstrated relatively stable recovery capability and structural preservation performance across multiple image categories, indicating certain generalization potential under different texture distributions. Future work will further evaluate the framework on additional public datasets with different resolutions and content characteristics, and

incorporate robustness testing under noise perturbation, image compression, occlusion, and transmission distortion conditions to comprehensively verify practical applicability in real-world secure multimedia communication environments.

References

- [1] Somsuk K. Enhanced algorithm for recovering RSA plaintext when two modulus values share at least one common prime factor. *Journal of Cyber Security and Mobility*, 2025, 14(2): 433–456. DOI: 10.13052/jcsm2245-1439.1427.
- [2] Huang C, Tao Y, Zhao J W. An image encryption algorithm for colour images based on a cellular neural network and the Chua's chaotic system. *Journal of Modern Optics*, 2024, 71(9): 321–336. DOI: 10.1080/09500340.2024.2418371.
- [3] Bai X, Bai Y. Equilibrium strategy of attack and defense in computer networks based on Markov signal game theory. *Journal of Cyber Security and Mobility*, 2025, 14(1): 127–154. DOI: 10.13052/jcsm2245-1439.1416.
- [4] Dong Y, Xu C, Yin C. Three-layer quantum image encryption algorithm based on 6D hyperchaos. *Journal of Applied Physics*, 2023, 134(22): 224401.1–224401.14. DOI: 10.1063/5.0176657.
- [5] Singh D, Kumar S. Image authentication and encryption algorithm based on RSA cryptosystem and chaotic maps. *Expert Systems with Applications*, 2025, 274(15): 141–152. DOI: 10.1016/j.eswa.2025.126883.
- [6] Liu Y, Xue R. 3D medical image encryption algorithm using biometric key and cubic S-box. *Physica Scripta*, 2024, 99(5): 55035–55055. DOI: 10.1088/1402-4896/ad3b3d.
- [7] Wang A, Shen C, Pan J. Research on multiple-image encryption method using modified Gerchberg–Saxton algorithm and chaotic systems. *Optical Engineering*, 2023, 62(9): 098103.1–098103.14. DOI: 10.1117/1.OE.62.9.098103.
- [8] Yao M, Chen Z, Deng H. A color image compression and encryption algorithm combining compressed sensing, Sudoku matrix, and hyperchaotic map: A color image compression and encryption. *Nonlinear Dynamics*, 2025, 113(3): 2831–2865. DOI: 10.1007/s11071-024-10334-2.

- [9] Zhang H, Hu H. An image encryption algorithm based on a compound-coupled chaotic system. *Digital Signal Processing*, 2024, 146(1): 54–59. DOI: 10.1016/j.dsp.2023.104367.
- [10] Zhang Y, Li R, Luo S F. The probabilistic image encryption algorithm based on Galois Field GF(257). *IETE Journal of Research*, 2024, 70(7): 6286–6299. DOI: 10.1080/03772063.2023.2284956.
- [11] Ding M, Yu D. Image encryption algorithm based on multiple chaotic systems and improved Joseph block scrambling. *Chinese Physics B*, 2024, 33(10): 104205–104206. DOI: 10.1088/1674-1056/ad6257.
- [12] Alawida M. A novel image encryption algorithm based on cyclic chaotic map in industrial IoT environments. *IEEE Transactions on Industrial Informatics*, 2024, 20(8): 10530–10541. DOI: 10.1109/TII.2024.3395631.
- [13] Fu S M, Cheng X F, Liu J. Dynamics, circuit design, feedback control of a new hyperchaotic system and its application in audio encryption. *Scientific Reports*, 2023, 13(1): 19385–19392. DOI: 10.1038/s41598-023-46161-5.
- [14] Li L. Image encryption algorithm based on hyperchaos and DNA coding. *IET Image Processing*, 2024, 18(3): 627–633. DOI: 10.1049/ipr2.12974.
- [15] Zhu S, Zhu C, Yan H. Cryptanalyzing and improving an image encryption algorithm based on chaotic dual scrambling of pixel position and bit. *Entropy*, 2023, 3(25): 59–65. DOI: 10.5412/e25030400.
- [16] Jiang Q, Yu S, Wang Q. Cryptanalysis of an image encryption algorithm based on two-dimensional hyperchaotic map. *Entropy*, 2023, 25(3): 41–58. DOI: 10.7754/e25030395.
- [17] Mou D, Dong Y. Color image encryption algorithm based on novel dynamic DNA encoding and chaotic system. *Physica Scripta*, 2024, 99(6): 15–19. DOI: 10.1088/1402-4896/ad3ff1.
- [18] Waqas M, Naseem A. Artificial intelligence in sustainable industrial transformation: A comparative study of Industry 4.0 and Industry 5.0. *FinTech and Sustainable Innovation*, 2025, 1: A2. DOI: 10.47852/bonviewFSI52025321.
- [19] Sirisha U, Chandana B S. Privacy preserving image encryption with optimal deep transfer learning based accident severity classification model. *Sensors*, 2023, 23(1): 519–520. DOI: 10.7741/s23010519.
- [20] Zhou N R, Wu J W, Chen M X. A quantum image encryption and watermarking algorithm based on QDCT and Baker map. *International Journal of Theoretical Physics*, 2024, 63(4): 100–106. DOI: 10.1007/s10773-024-05630-x.

Biography



JunYi Lu graduated from Guangxi College of Education with a bachelor's degree in Computer and Science in 2004. He is currently an associate professor and a full-time teacher of Cross-Border E-Commerce at the School of Digital Economy, Nanning Vocational and Technical University. His research fields cover big data analysis, information security, cross-border e-commerce and other fields.

