Journal of Artificial Intelligence and Capsule Networks (ISSN: 2582-2012)
www.irojournals.com/aicn/

Reversible Data Hiding in Encrypted Images

Raagul P.[1], Subash P R.[2], Deepak Indirajith M C.[3], Sathya T.[4]

[1,2,3]Student, [4]Assistant Professor, Department of Artificial Intelligence and Data Science, NPR College of Engineering and Technology, Dindigul, India.

E-mail: [1]raagulp920822243041@nprcolleges.org, [2]subashpr920822243052@nprcolleges.org, [3]deepakindirajithmc920822243015@nprcolleges.org, [4]sathyasethuezhil@gmail.com

Abstract

The growing demand for secure digital communication has led to the necessity of developing strong schemes that can provide confidentiality, integrity, and information recovery. In general, traditional image encryption and steganography methods suffer from limitations due to the reduction in either embedding capacity or visual quality. In this paper, a novel scheme for secure reversible data hiding of an encrypted image based on the combination of Password-Based Key Derivation Function 2 (PBKDF2), Advanced Encryption Standard in Galois/Counter Mode (AES-GCM), and Most Significant Bit (MSB) embedding is introduced. In the first place, PBKDF2 uses a password chosen by the user to derive a secret key, which is used in order to encrypt the secret message using AES-GCM. Then, the generated ciphertext is embedded into the cover image via MSB substitution method. The experimental analysis performed with the help of USC-SIPI image database shows that the suggested scheme provides 52.36 dB PSNR, 0.45 MSE, 0.99 SSIM and 2.84 bpp embedding capacity.

Keyword: Reversible Data Hiding in Encrypted Images (RDH-EI), Authenticated Image Encryption, Password-Based Key Derivation Function (PBKDF2), AES-GCM Cryptographic Security, Imperceptible Data Embedding.

1. Introduction

The fast growth of digital communication, cloud computing, and multimedia data exchange has greatly contributed to the necessity of creating some kind of security measure

DOI: https://doi.org/10.36548/jaicn.2026.3.007
Extended version of article received from International Conference on Smart Innovations in Engineering & Technology (ICSIET 2026)
Received: 28.05.2026, received in revised form: 30.06.2026, accepted: 18.07.2026, published: 29.07.2026

that will ensure the privacy of the transmitted data. At present, images are among the most commonly used media for transferring confidential data in healthcare, military, finance, legal, and governmental sectors. While traditional cryptographic methods successfully protect the content of the information being transferred, they do not hide the very existence of the encrypted data and therefore make it possible to intercept, analyze and attack. In a similar way, transmitting confidential data via regular channels poses the risk of data breach and violation of data integrity. This issue is what brought about the need for creating some kind of secure communication system which would provide confidentiality, authenticity, and imperceptible transfer of data. Therefore, combining the process of encryption and data hiding has proved to be an effective solution to this problem.

There is an ongoing effort in the existing literature where several cryptographic techniques, steganography and RDH techniques have been examined to ensure security of information. The latest cryptographic techniques have proved to provide high level security from any form of illegal intrusion. Similarly, hybrid cryptographic models have provided security in the field of communication by using various encryption techniques [1], [2]. In the same way, secret data can be embedded in the encrypted images by using reversible data hiding techniques, allowing recovery of secret data along with the recovery of original images in their initial state [3], [4]. Moreover, latest research has focused on embedding capacity and image quality through different techniques [5], [6], [10]. In spite of all these developments, most of the current techniques treat encryption and data embedding independently. Furthermore, concepts like authenticated encryption, strong key generation, reversible embedding, and light-weight web implementation have not been put together to form a coherent framework. These gaps in the field make the development of a scalable and secure technique essential.

To solve these challenges, DataCrypt is introduced as a solution, which is a secure framework for reversible data hiding in encrypted images. The framework utilizes a combination of authenticated encryption and reversible image steganography in order to enable reliable and covert communication. The framework utilizes PBKDF2-based key generation, AES-GCM encryption, and MSB-based reversible data embedding in a web-based environment. The framework provides an improvement on confidentiality, integrity, and communication privacy by combining robust cryptographic techniques with an invisible embedding technique that does not change the transmission process. The framework is intended to provide secure communication and reversible data hiding by supporting the transmission of

messages via user authentication. The performance of the proposed framework is analyzed experimentally in order to demonstrate its effectiveness in providing secure communication and data hiding.

2. Literature Review

The growing necessity of secure digital communication has brought advancements in the field of encryption and reversible data hiding techniques for protecting private information during transmission processes. The new generation of encryption algorithms of images has succeeded in enhancing the privacy of the transmitted information using effective encryption techniques to facilitate safe communication with minimal computation cost [1]. The hybrid cryptography schemes using block cipher techniques have also contributed significantly to enhancing the security of communications by combining different encryption techniques to resist different cyberattacks [2]. Although encryption has shown its efficiency, the encrypted data is visually detectable.

In recent years, RDH-EI has become an efficient approach that not only guarantees the confidentiality of the information but also helps recover the original image perfectly once the information is extracted from it. Multi-MSB compression techniques for embedding have enhanced the embedding efficiency while keeping the image quality intact and ensuring reversibility [3]. High-capacity multi-layer embedding schemes have also helped increase the embedding capacity without affecting the image quality much; therefore, these techniques have been successfully used in secure image communication systems [4].

Modern researches have paid attention to improving the effectiveness of embedding through adaptive and prediction-based approaches. Odd-even discrimination approach has increased the effectiveness of embedding by using proper places for embedding, thus minimizing the distortion in images [7]. Entropy encoding approach and pixel prediction-based approach have managed to increase payload while maintaining high reconstruction quality and extraction accuracy [9]. In addition, separable reversible data hiding schemes have provided an advantage of being able to extract data independently from each other and recover images [10]. Moreover, pixel-difference-based room reservation technique has proved more efficient in embedding as well as in lossless extraction of encrypted images [8]. However, even though there have been many improvements in the field of reversible data hiding, very few researches have considered the usage of authenticated encryption, robust key generation and web-based

implementation. These gaps have led to the development of DataCrypt framework, which utilizes key generation through PBKDF2, AES-GCM encryption and MSB-based reversible data hiding technique.

3. Proposed Methodology

The proposed framework is created to ensure secure and reversible data embedding in an image by utilizing authenticated encryption alongside image steganography techniques. Thus, the process of creation of cryptographic keys, data encryption with AES-GCM algorithm and data embedding by means of MSBs will be integrated into one flow of operations to provide confidentiality, integrity and imperceptible communication. At first, authenticated users supply a secret message and a cover image for encryption and embedding. After that, the cryptographic key is generated by means of PBKDF2, and then the message is encrypted with AES-GCM into authenticated ciphertext. The ciphertext is then embedded into the chosen cover image via MSB substitution to create a stego image with minimal visible changes. In case of extraction, the embedded ciphertext is retrieved and decrypted to get back the initial message. The sequence of the operations within the proposed framework is shown in Figure 1.

3.1 Framework Overview

This proposed system implements an algorithm that involves sequential processing that incorporates the process of authentication, encryption using cryptographic techniques, and reversible data hiding to allow secure image-based communication. The whole process starts with the authentication process, where secret message collection and cover image selection is done. The generation of a secure cryptographic key is done by using the PBKDF2 key derivation function that is used by the AES-GCM encryption technique to encrypt the input message and at the same time maintaining its integrity and security. Finally, this message is embedded into the cover image using the MSB-based embedding technique to generate a secure and imperceptible stego image. On the receiver side, this embedded information is recovered from the stego image and is authenticated and decrypted using the same cryptographic key.

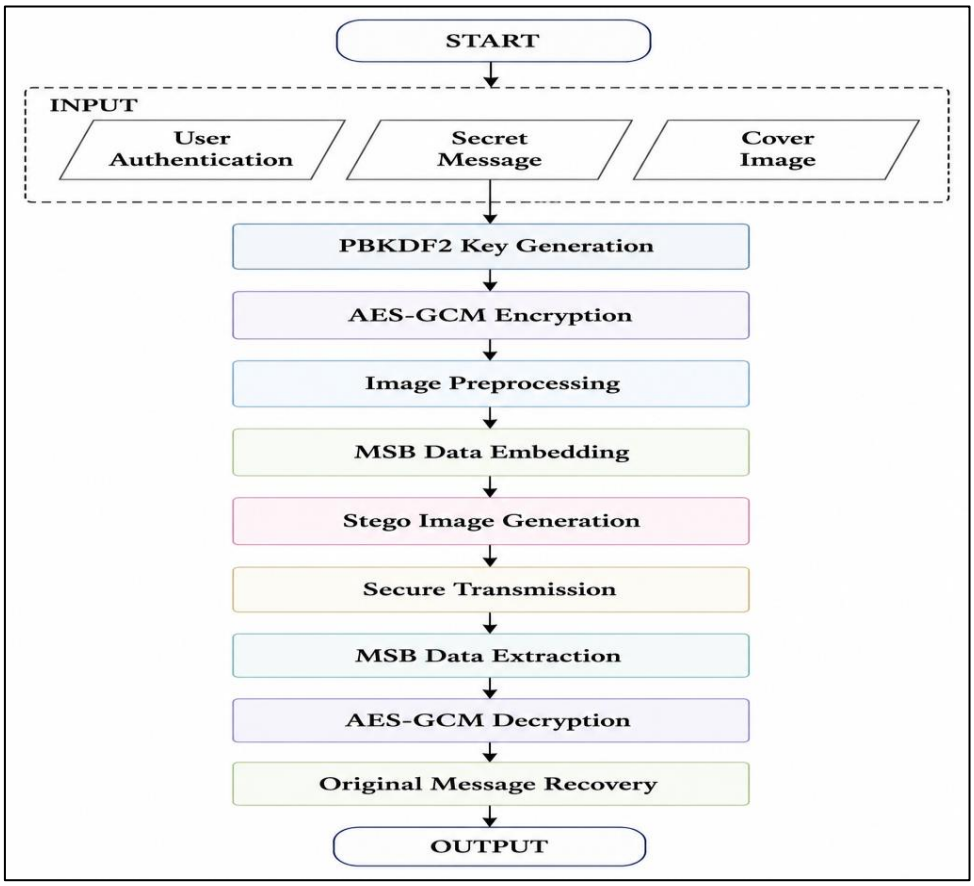

Figure 1. Workflow of the proposed reversible data hiding framework

Figure 1 shows the entire process involved in the proposed framework. The framework is built around secure user authentication, encryption and decryption, embedding process, secure transmission of embedded information, and message retrieval. The above mentioned process ensures confidential communication without compromising on the integrity of data and accuracy of message recovery at the receiver's end.

3.2 Dataset Description

The experimental testing of the developed system used test images sourced from the USC-SIPI Image Database [11] and commonly used for research purposes in image processing, watermarking, steganography, and data hiding. The database includes a wide range of gray scale and color images having varied structures, thereby facilitating comprehensive experiments for varied image structures. For the consistency of experimental results, a set of representative images of 256 × 256, 512 × 512, and 1024 × 1024 pixels were chosen to study the effect of image size on encryption, embedding capability, and data retrieval.

Unlike other methods based on machine learning, the suggested approach does not need training/testing data sets. Rather, text messages defined by the user are generated dynamically and encrypted before embedding into the selected cover images. Prior to the embedding process, the images are transformed into their RGB pixels form and preprocessed to embed data using MSBs. The features of the data set and input parameters of the experiment are listed in Table 1.

Table 1. Characteristics of the USC-SIPI image database

Parameter	Specification
Dataset Name	USC-SIPI Image Database [11]
Total Images	632 images
Image Categories	Textures, Aerials, Miscellaneous, and Sequences
Texture Images	154
Aerial Images	38
Miscellaneous Images	44
Sequence Images	396 (69 image sequences)
Image Type	Color and Grayscale
Image Resolution	256×256, 512×512, 1024×1024 pixels (varies by category)
Image Format	TIFF (original), converted to PNG/JPEG for experimentation

3.3 Secure Data Hiding Process

In the suggested scheme, a mixed approach is used for the purpose of ensuring that confidential and reliable image communication is accomplished through the use of a combination of authenticated encryption and reversible data hiding. In the first place, a cryptographic key is generated out of the user-defined password via the use of PBKDF2. The key thus generated will be used to encrypt the secret message using the AES-GCM, which at the same time generates an authentication tag. The mathematical representation of the key derivation process is

$$K = \text{PBKDF2}(P, S, c, L) \tag{1}$$

where P denotes the user password, S represents the randomly generated salt, c is the iteration count, and L is the required key length.

Following the secure key generation, the plaintext is encoded into ciphertext using the AES-GCM encryption scheme. This encryption scheme ensures security by combining confidentiality and integrity, which means that the data cannot be tampered with in transit. The process of authenticated encryption can be described by

$$(C, T) = \text{AES} - \text{GCM}_K(N, M, A) \tag{2}$$

where M is the plaintext message, K is the derived encryption key, N denotes the initialization nonce, A represents the additional authenticated data, C is the encrypted ciphertext, and T is the authentication tag. The entire process of encryption implemented in the suggested system is shown in Figure 2.

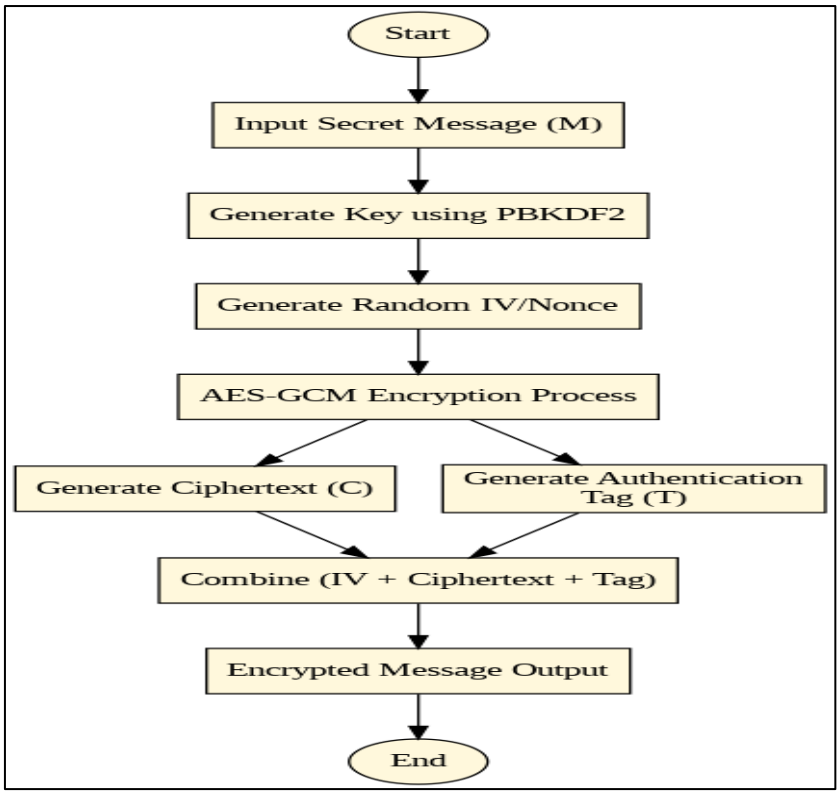

Figure 2. AES-GCM encryption process

Figure 2 shows the step of authenticated encryption in the proposed scheme. The combination of AES-GCM guarantees that the message will be both confidential and authentic before the encrypted message is embedded in the host image.

After encryption, the chosen cover image undergoes preprocessing, which involves changing pixel values to binary format for manipulation at the bit level. This process is followed by embedding the encrypted binary code in the image via the Most Significant Bit (MSB) substitution method, resulting in a stego-image without altering the features of the cover image. The embedding process can be shown as

$$I_s = (I_c \wedge M_b) \mid D_b \tag{3}$$

where I_c denotes the cover image pixel, M_b is the masking operation, D_b represents the encrypted data bit, and I_s is the resulting stego image. The embedding procedure is depicted in Figure 3.

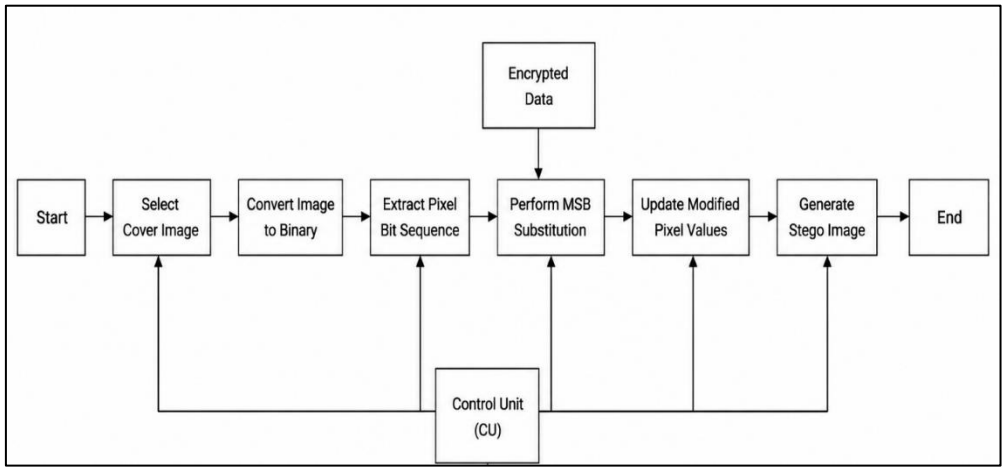

Figure 3. MSB-based data embedding process.

Figure 3 shows the technique used for embedding the encrypted message into the cover image. The bit substitution technique ensures that the data is embedded safely without compromising its visual quality.

At the receiving end, the information embedded in the stego image is retrieved through an inverse MSB process. To ensure the validity of the ciphertext before decryption, it goes through an authenticity check using the AES-GCM authentication tag. After successful authentication, the original plaintext is obtained without losing any information. The process of extraction can be represented as

$$D_b = I_s M_b \tag{4}$$

where D_b denotes the extracted encrypted data bits, I_s represents the stego image, and M_b is the extraction mask. The complete extraction and recovery procedure are illustrated in Figure 4.

The inverse extraction procedure successfully extracts the hidden ciphertext from the stego-image without disturbing the characteristics of the original image. The extracted ciphertext is verified and decrypted by using the appropriate decryption key, and hence the original secret message is successfully recovered. The reversible nature of the extraction technique guarantees secure and efficient transmission of messages without altering the characteristics of the cover image.

3.4 Performance Evaluation

The effectiveness of the proposed framework is tested based on quantifiable image quality and embedding performance measures. The analysis mainly concentrates on determining the level of visual distortion, ability to embed information, and the ability to extract the information after encryption and embedding. Four key metrics are used, which include Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Embedding Capacity (EC), and Extraction Accuracy (EA).

The MSE calculates the average distance between pixels in the cover image and the generated stego image using the following equation 5:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j) - S(i,j)]^2 \tag{5}$$

where $I(i,j)$ and $S(i,j)$ denote the pixel intensities of the original and stego images, respectively.

The PSNR estimates the visual quality of the generated stego image using the following equation 6:

$$\text{PSNR} = 10\log_{10}\left(\frac{MAX^2}{\text{MSE}}\right) \tag{6}$$

where MAX represents the maximum possible pixel value of the image.

The embedding capability of the proposed framework is quantified using Embedding Capacity (EC), expressed in bits per pixel (bpp), as

$$\mathrm{EC} = \frac{B}{M \times N} \tag{7}$$

where B denotes the total number of embedded bits, while M and N represent the image dimensions.

The accuracy of the recovery process is evaluated using Extraction Accuracy (EA), which is defined as

$$\mathrm{EA} = \frac{N_c}{N_t} \times 100 \tag{8}$$

where N_c is the number of correctly recovered bits and N_t is the total number of embedded bits.

It should be noted that the set of criteria mentioned above gives the possibility to assess the developed system in terms of three important aspects, namely image quality, embedding capacity, and extraction accuracy. The MSE and PSNR criteria are applied to estimate the quality of stego images produced in the course of the work. Embedding Capacity criterion is used to determine the quantity of encrypted messages that can be inserted into the image. Extraction Accuracy is used to verify whether the message extracted from the image was decrypted correctly.

3.5 Experimental Setup.

The proposed framework is implemented using Python 3.11 programming language in combination with Flask web framework for the development of the secure communication interface. The experimental implementation was done in a computer with Intel Core i7 Processor and 16GB RAM running Windows 11 operating system. The framework implements PBKDF2 key derivation algorithm, AES-GCM authenticated encryption, and MSB reversible data hiding algorithm to test secure embedding and extraction of message. The benchmark cover images used in the experiment were from the USC-SIPI Image Database, and text messages specified by the user were dynamically encrypted and embedded into these cover images during the experiment. The experimental performance has been evaluated through the

following metrics such as PSNR, MSE, Embedding Capacity, and Extraction Accuracy. The proposed experimental configuration details are summarized in Table 2.

Table 2. Experimental configuration

Parameter	Specification
Development Language	Python 3.11
Framework	Flask
Database	MySQL
Operating System	Windows 11
Processor	Intel Core i7
Memory	16 GB RAM
Image Formats	PNG, JPEG
Encryption Algorithm	AES-GCM (256-bit)
Data Hiding Technique	MSB Substitution

4. Results and Discussion

The effectiveness of the proposed framework was assessed experimentally to evaluate its suitability for the purpose of secure reversible data hiding by maintaining the image quality and providing for secure recovery of the message. This was done through the analysis of experimental results based on the use of objective measurements of the image quality together with subjective evaluation of the stego images created and the messages extracted. Performance assessment will be carried out based on the evaluation of the efficiency of the encryption process, embedding ability, image fidelity, and extraction accuracy under varied experimental conditions.

4.1 Performance Analysis

Quantitative evaluation of the proposed framework was done using some image quality and embedding measures such as Encryption Time, PSNR, MSE, Embedding Capacity, and Extraction Accuracy. These measures enable an overall evaluation of the computational efficiency, visual quality, embedding capability, and reliability of the reversible data hiding

process. The results of the experiment conducted with different test images are given in Table 3.

Table 3. Performance evaluation of the proposed framework

Performance Metric	**Observed Value**	**Performance Objective**
Encryption Time (ms)	18.60	Fast secure encryption
PSNR (dB)	52.36	High image fidelity
MSE	0.45	Low embedding distortion
SSIM	0.99	Structural similarity preservation
Embedding Capacity (bpp)	2.84	Efficient payload embedding
Extraction Accuracy (%)	100.0	Accurate message recovery

The quantitative analysis illustrated in Table 3 confirms the efficiency of the suggested approach to securing the image and allowing reversible data hiding without significant loss of image quality. The fact that the PSNR and SSIM indices are very high is the indication that there is almost no visible distortion, while the very low MSE value proves that the information from pixels is preserved precisely. Also, the obtained embedding capacity allows concealing the encrypted data without any effect on image quality. Perfect extraction of the message embedded into the image, as confirmed by extraction accuracy, proves the efficiency of PBKDF2, AES-GCM, and MSB-based embedding approach.

4.2 Visual Performance Analysis

To test the visual aspect of the generated stego image, the latter was compared with its respective original cover image post the embedding process. Apart from visual assessment, the difference image was inspected to determine how many pixel changes were made in reversible data hiding. This gives us the proof that the embedding process does not affect the perceptual properties of the cover image but at the same time hides the encryption. Figure 4 depicts the resultant visual output.

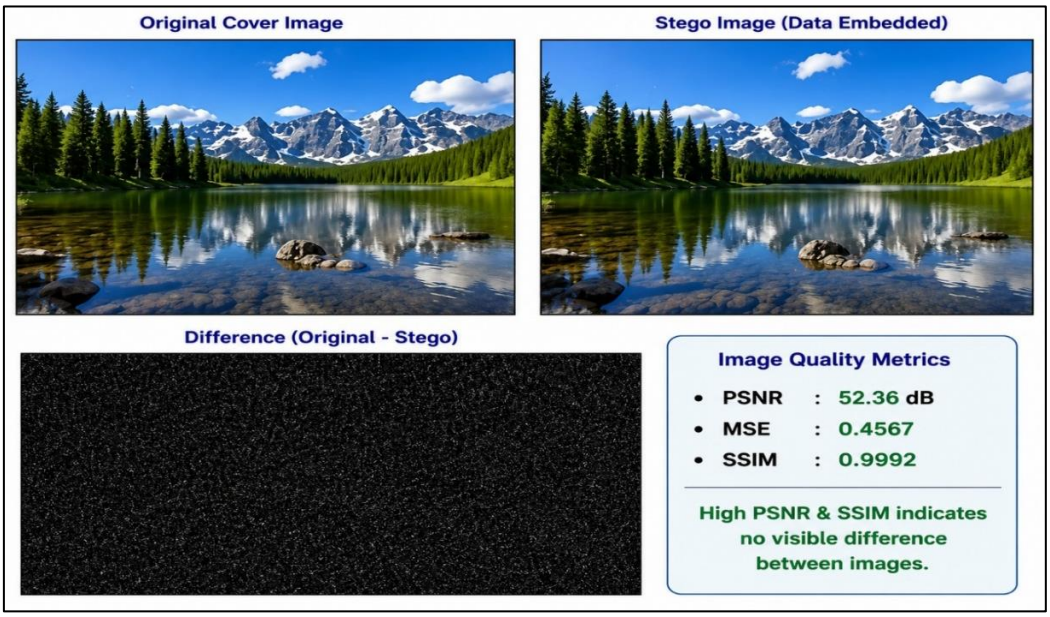

Figure 4. Visual evaluation of the stego image generated by the proposed framework.

It can be observed from Figure 4 that even after successful embedding of the encrypted data, the stego image is visually very similar to the cover image used. The difference image indicates that there are only a few numbers of pixel values that change during the embedding process, which makes it clear that the MSB-based data hiding algorithm presented is imperceptible. Moreover, the results regarding the image quality measures show that there was no distortion introduced in the embedding procedure.

4.3 Embedding Quality Analysis

The embedding quality of the proposed method was investigated through examining the variation of image quality with varying capacities of embedding. Due to the fact that reversible data hiding involves modification of pixels values in order to include the encrypted messages, it is essential to investigate the relation between the payload size and the quality of the image. The relation between PSNR and embedding capacity becomes a good indicator for this purpose. The resultant PSNR variations are shown in Figure 5.

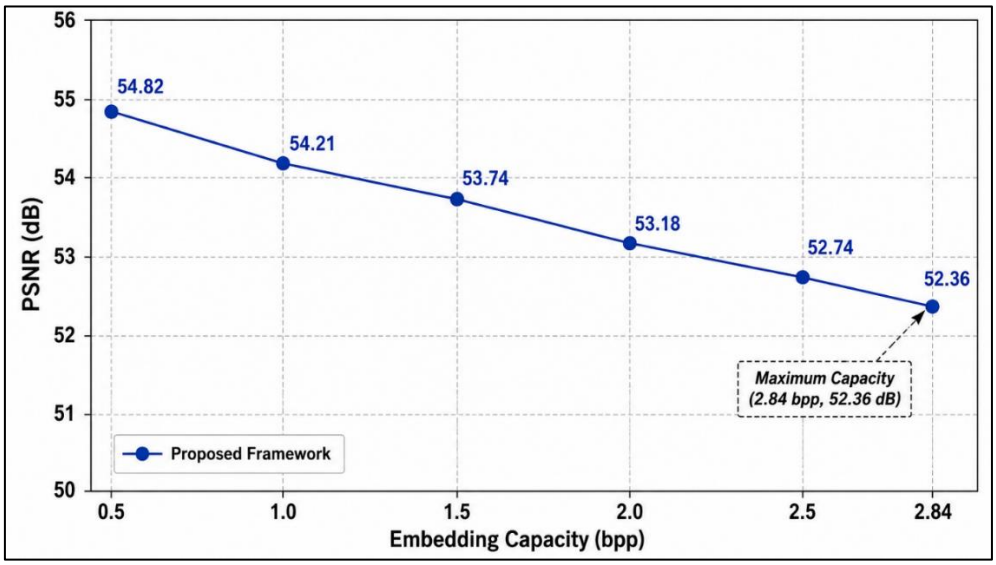

Figure 5. PSNR variation with respect to embedding capacity.

From Fig. 5, it is clear that PSNR value decreases slowly as the embedding capacity increases due to the reason that more payload will involve more changes in the pixels while embedding. Although the values of PSNR have decreased, yet these values are still greater than 52 dB, thus ensuring that the resultant images have very good quality. This shows that the PBKDF2-AES-GCM-MSB scheme can accommodate large-sized payloads in an efficient manner. Therefore, it can be said that the proposed scheme of reversible data hiding is quite efficient and effective.

The results of the experiments collectively show that the suggested scheme serves as a good solution for secure reversible data hiding in encrypted images using PBKDF2-based key generation, AES-GCM authenticated encryption, and MSB-based embedding. The analysis of the quality of the stego images proves that the images generated have very good perceptual quality, and the analysis of the quality of the embedding shows that the increase in the size of the embedded data has a slight influence on the quality of the image. Moreover, the scheme guarantees that the embedded data can be reliably recovered through authenticated decryption.

5. Conclusion and Future Work

A reversible data hiding method for encrypted images was discussed in this study by employing PBKDF2 algorithm based key generation technique, AES-GCM authenticated encryption technique and MSB-based data hiding. The proposed approach allows for secure message transmission in addition to the quality assurance of images and the complete recovery

of hidden information without any loss. According to experimental analysis, the proposed framework has maintained high visual quality as indicated from PSNR of 52.36 dB, MSE of 0.45, SSIM of 0.99, embedding rate of 2.84 bpp, and 100% extraction accuracy. This has proven the capability of the approach in achieving the security and invisibility of data hiding. The proposed framework can be applied for multimedia communication, image sharing, and any other application where security and image preservation are required. Future researches will be conducted for the development of the framework for supporting color and high-resolution images with adaptive data hiding schemes that ensure high payload capacity and image quality. Furthermore, lightweight optimization and post-quantum cryptographic mechanisms will be adopted for the enhancement of efficiency and security of future communications systems.

References

[1] Raman, Ramakrishnan, Yassir Farooqui, K. Nisha, and Manisha Paliwal. "Efficient Encryption Techniques for Secure Transmission of Dynamic Image." In 2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS), IEEE, 2025: 1-6.

[2] Gour, Ayush, Simarjit Singh Malhi, Gursharan Singh, and Gursimran Kaur. "Hybrid Cryptographic Approach: for Secure Data Communication Using Block Cipher Techniques." In E3S Web of Conferences, EDP Sciences, 2024, vol. 556: 01048.

[3] Chen, Fan, Yuan Yuan, Hongjie He, Miao Tian, and Heng-Ming Tai. "Multi-MSB Compression Based Reversible Data Hiding Scheme in Encrypted Images." IEEE Transactions on Circuits and Systems for Video Technology 2020, vol 31, no. 3: 905-916.

[4] Malik, Asad, Peisong He, Hongxia Wang, Ahmad Neyaz Khan, Saied Pirasteh, and Sani M. Abdullahi. "High-Capacity Reversible Data Hiding in Encrypted Images Using Multi-Layer Embedding." IEEE Access 2020, vol 8: 148997-149010.

[5] Stallings, William. "Cryptography and Network Security: Principles and Practice 2017, Global Edition: 7.

[6] Chen, Chih-Cheng, Chin-Chen Chang, and Kaimeng Chen. "High-Capacity Reversible Data Hiding in Encrypted Image Based on Huffman Coding and Differences of High

Nibbles of Pixels." Journal of Visual Communication and Image Representation 2021, vol 76: 103060.

[7] Agarwal, Ruchi, Ankita Vaish, and Manoj Kumar. "Odd-Even Discrimination-Based Reversible Data Hiding Technique in Encrypted Images." Journal of Electronic Imaging 2023, vol 32, no. 1: 013032-013032.

[8] Venkatesh, Veeramuthu, R. Anushiadevi, Padmapriya Velupillai Meikandan, Hemalatha Mahalingam, and Rengarajan Amirtharajan. "Separable Reversible Data Hiding by Vacating Room After Encryption Using Encrypted Pixel Difference." Scientific Reports 2025, vol 15, no. 1: 11916.

[9] Qiu, Yingqiang, Qichao Ying, Yuyan Yang, Huanqiang Zeng, Sheng Li, and Zhenxing Qian. "High-Capacity Framework for Reversible Data Hiding in Encrypted Image Using Pixel Prediction and Entropy Encoding." IEEE Transactions on Circuits and Systems for Video Technology 2022, vol 32, no. 9: 5874-5887.

[10] Anushiadevi, R., Padmapriya Praveenkumar, John Bosco Balaguru Rayappan, and Rengarajan Amirtharajan. "Uncover The Cover to Recover the Hidden Secret-A Separable Reversible Data Hiding Framework." Multimedia Tools and Applications 2021, vol 80, no. 13: 19695-19714.

[11] USC Signal and Image Processing Institute, The USC-SIPI Image Database. University of Southern California. Available: https://sipi.usc.edu/database/