ISSN 2277-2685
IJESR/Jan-Mar. 2026/ Vol-16/Issue-1/181-187
Mr. Mirza Sajid Ali Baig *et. al.,* / International Journal of Engineering & Science Research

Reversible Data Hiding In Images Using Dual Embedding And Pixel Prediction

Mr. Mirza Sajid Ali Baig[1], Aledhandi Sheshank[2], Sachin Bemanavena[3], A. Mohammed Wajid[4], Chandrakanti Likith Rao[5]

[1]Assistant Professor, Department of Electronics and Communication Engineering, Teegala Krishna Reddy Engineering College, Hyderabad, India.
[2,3,4,5]B.Tech Students, Department of Electronics and Communication Engineering, Teegala Krishna Reddy Engineering College, Hyderabad, India.

Mirzaece@tkrec.ac.in, sheshankaledhandi@gmail.com, Sachinbeemanavena@gmail.com, amohammedwajid30@gmail.com, Chandrakantilikithrao@gmail.com

ABSTRACT

Reversible Data Hiding (RDH) has become an essential research area in digital image security because it allows confidential information to be embedded into digital images while ensuring perfect recovery of both the hidden data and the original cover image. Unlike traditional steganography and watermarking techniques that introduce permanent distortion, RDH guarantees complete reversibility, making it highly suitable for sensitive domains such as medical imaging, military communication, forensic analysis, and digital archiving. Over the years, various RDH techniques such as Difference Expansion, Histogaaram Shifting, and Prediction Error Expansion have been proposed to enhance embedding capacity and minimize distortion [14], [6], [15]. However, achieving an optimal balance between embedding capacity, visual fidelity, and computational simplicity remains a challenge.

This research proposes a dual embedding framework that integrates Prediction Error Expansion (PEE) with Least Significant Bit (LSB) substitution to improve embedding efficiency while preserving high image quality. Pixel prediction using a Median Edge Detector (MED) predictor is employed to reduce prediction errors and minimize embedding distortion. The proposed method is evaluated using Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Structural Similarity Index Measure (SSIM). Experimental results demonstrate that the proposed technique achieves PSNR values above 57 dB even at higher payload capacities, outperforming traditional histogram shifting and standalone LSB approaches [1], [6], [14]. The findings confirm that the proposed dual embedding approach effectively balances capacity, imperceptibility, and reversibility.

Keywords: *Reversible Data Hiding, Prediction Error Expansion, Least Significant Bit, Pixel Prediction, Image Security, PSNR, SSIM, Data Embedding.*

1. INTRODUCTION

The rapid advancement of digital communication technologies has significantly increased the need for secure transmission and storage of confidential information. Digital images are widely used in social media, medical imaging, defense communication, and digital documentation, making them ideal carriers for hidden data transmission. Reversible Data Hiding (RDH) is a specialized branch of information hiding that allows embedding secret information into a digital image while ensuring exact restoration of the original image after extraction of the hidden data. This distinguishing characteristic makes RDH particularly valuable in sensitive applications where any permanent distortion is unacceptable [15].

Early RDH techniques were based on lossless compression strategies. The R–S scheme introduced by Fridrich and colleagues utilized compressed bit planes to create embedding space [3]. Later, Difference Expansion (DE) was introduced to increase embedding capacity by expanding the difference between adjacent pixels [14]. Although DE improved payload capacity, it sometimes resulted in noticeable visual distortion when embedding larger amounts of data.

Histogram Shifting (HS) methods were subsequently proposed to minimize distortion by modifying histogram peak points and shifting neighboring bins to accommodate secret bits [6]. While HS significantly reduced distortion, it often suffered from limited embedding capacity. Prediction-based RDH methods then emerged as a more efficient solution. Prediction Error Expansion (PEE), introduced by Thodi and Rodríguez [14], utilized prediction errors instead of raw pixel differences, thereby achieving improved fidelity and embedding efficiency.

Recent developments have focused on enhancing pixel prediction accuracy using adaptive predictors and pixel value ordering techniques [11], [12]. These methods reduce prediction errors, leading to smaller embedding distortion. Despite these improvements, simple techniques such as LSB substitution remain popular due to their simplicity and high embedding capacity [2]. However, standalone LSB lacks robust reversibility guarantees.

Motivated by the need to balance simplicity, capacity, and fidelity, this research integrates PEE with LSB substitution in a dual embedding framework to improve overall performance while maintaining computational efficiency.

2. LITERATURE REVIEW

Reversible Data Hiding (RDH) has evolved significantly over the past two decades, with research primarily focusing on improving embedding capacity while preserving image fidelity and guaranteeing lossless recovery. Early foundational work by Fridrich et al. introduced the R–S scheme, which utilized lossless compression of bit planes to create embedding space in images [3]. Although effective, the compression-based strategy was limited by the redundancy present in the cover image and often resulted in constrained payload capacity.

A major breakthrough occurred when Tian introduced the Difference Expansion (DE) technique, which expanded the difference between neighboring pixels to embed secret data [14]. DE substantially improved embedding capacity compared to earlier compression-based methods. However, at higher payloads, noticeable distortion was introduced, especially in smooth image regions.

To address distortion control issues, Tai et al. proposed histogram modification of pixel differences, shifting histogram bins to embed data while maintaining lower perceptual impact [7]. Histogram Shifting (HS) techniques became popular due to their simplicity and reduced distortion, though embedding capacity remained relatively limited.

To enhance embedding efficiency further, prediction-based RDH methods were developed. Tseng et al. introduced a prediction-based reversible data hiding approach combining lossless compression and group classification strategies, improving embedding performance compared to pure DE methods [3]. Thodi and Rodríguez later proposed Prediction Error Expansion (PEE), which replaced raw pixel differences with prediction errors computed from neighboring pixels [14]. Because prediction errors are typically smaller than raw pixel differences, PEE reduces embedding distortion and improves visual fidelity.

Hong et al. further refined prediction-based embedding by modifying prediction errors adaptively to achieve higher image quality [5]. Their work demonstrated that accurate prediction directly influences embedding performance. Similarly, Kumar et al. provided a comprehensive review of prediction techniques in RDH and concluded that adaptive predictors outperform fixed predictors in achieving better PSNR values and embedding efficiency [1].

Jung et al. explored LSB-based data hiding combined with image interpolation techniques to improve payload efficiency while maintaining reversibility [2]. Although LSB substitution offers high embedding capacity and low computational complexity, it lacks inherent distortion control and can become vulnerable to steganalysis if not carefully implemented.

Recent advancements have extended RDH into color and compressed domains. He and Cai proposed a reversible data hiding scheme for color images using channel reference mapping and adaptive pixel prediction, achieving PSNR values exceeding 60 dB at moderate payloads [6]. Their approach exploited inter-channel correlations to improve embedding fidelity. Mao et al. introduced channel unity embedding strategies that utilized cross-channel dependencies to reduce distortion in color images [10].

Qu and Kim proposed an enhanced pixel value ordering (PVO) predictor to reduce shifted pixel proportions during embedding, significantly improving fidelity at higher embedding rates [11]. Pixel value ordering techniques sort pixel intensities within blocks to exploit local correlations, reducing prediction errors and distortion. These improvements demonstrate that predictor accuracy plays a central role in RDH performance.

Hou et al. extended reversible data hiding into JPEG compressed images by modifying DCT coefficients and using block selection mechanisms [4]. Although compressed-domain RDH improves practical applicability, challenges such as bitstream expansion and synchronization errors remain. Zhang et al. presented a comprehensive survey summarizing two decades of RDH advancements and emphasized that prediction-error-based embedding remains the most promising direction for achieving high capacity and low distortion simultaneously [15].

More recent studies focus on encrypted-domain RDH, where data is embedded after image encryption to ensure dual-layer security. These methods enable secure cloud storage and privacy-preserving communication, though computational complexity increases significantly [15]. Additionally, adaptive embedding strategies based on local image complexity have been proposed to improve resistance against steganalysis attacks while maintaining reversibility [1].

From the literature, it is evident that no single method perfectly balances capacity, imperceptibility, computational efficiency, and reversibility. Prediction-based techniques offer superior fidelity, whereas LSB-based methods provide simplicity and high embedding rates. Inspired by these observations, the present research integrates Prediction Error Expansion with LSB substitution in a dual embedding framework to leverage the strengths of both approaches while minimizing their limitations.

3. METHODOLOGY

The methodology of the proposed reversible data hiding (RDH) system is designed to achieve a balanced trade-off between embedding capacity, visual fidelity, computational efficiency, and guaranteed reversibility. The framework integrates prediction-based embedding with Least Significant Bit substitution in a dual embedding architecture. This section describes the problem formulation, system design, block diagrams, and operational workflow in detail.

3.1 Problem Identification

Existing RDH techniques encounter a fundamental trade-off between embedding capacity and image quality. Increasing the payload typically increases the number of modified pixels, which may degrade visual fidelity and reduce Peak Signal-to-Noise Ratio (PSNR). While Difference Expansion methods provide relatively high embedding capacity, they may introduce visible artifacts at larger payloads [14]. Histogram Shifting techniques reduce distortion but often suffer from limited embedding capacity and require additional side information [7]. Prediction-based methods significantly reduce distortion by embedding data within prediction errors, but embedding flexibility may be constrained by prediction accuracy [1], [15]. Another challenge involves computational complexity. Advanced prediction models and encrypted-domain embedding techniques increase algorithmic complexity, limiting real-time implementation feasibility. Furthermore, many LSB-based methods provide high capacity but lack robust reversibility guarantees and may become vulnerable to steganalysis [2]. Therefore, a practical and efficient RDH method must simultaneously ensure:

- High embedding capacity
- Minimal perceptual distortion
- Low computational overhead
- Complete restoration of the original image

To address these limitations, the proposed work introduces a dual embedding framework combining Prediction Error Expansion (PEE) with LSB substitution.

3.2 Proposed System

The proposed system employs a hybrid embedding mechanism integrating pixel prediction, prediction error expansion, and selective LSB substitution. The cover image is first analyzed using a Median Edge Detector (MED) predictor. The MED predictor estimates the intensity of a target pixel based on its neighboring pixels, effectively preserving edge structures and reducing prediction error magnitude. Smaller prediction errors result in reduced distortion during embedding.

Once prediction values are computed, prediction errors are obtained by comparing actual and predicted pixel intensities. If the prediction error satisfies embedding conditions, secret bits are embedded using controlled expansion of the prediction error. For pixels that do not qualify for PEE embedding or when additional payload capacity is required, LSB substitution is applied in the blue channel. The blue channel is selected because the human visual system is less sensitive to intensity variations in this spectrum, thereby minimizing perceptual distortion [6], [14].

This dual mechanism ensures efficient utilization of pixel redundancy while maintaining high visual quality and complete reversibility.

3.3 Block Diagram Description

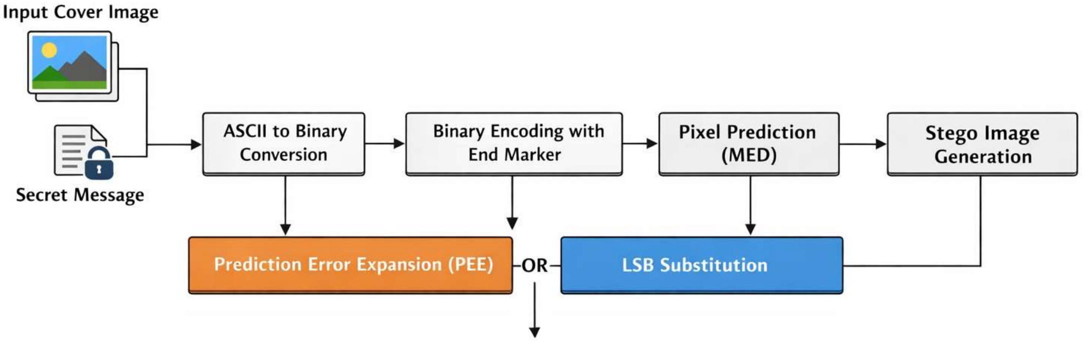

Figure 1: Block Diagram of Embedding Process

Figure 1 illustrates the architecture of the embedding stage. The process begins with the input of a cover image and a secret message. The secret message undergoes ASCII conversion followed by binary encoding. A predefined termination marker is appended to ensure proper extraction. Pixel prediction is performed using the MED predictor, and prediction errors are computed. Based on embedding conditions, either Prediction Error Expansion or LSB substitution is applied. Finally, modified pixel values are recombined to generate the stego image.

Flow Sequence in Figure 1:

Input Cover Image
→ Secret Message Conversion
→ Binary Encoding with End Marker
→ Pixel Prediction (MED)
→ Prediction Error Computation
→ Dual Embedding (PEE + LSB)
→ Stego Image Generation

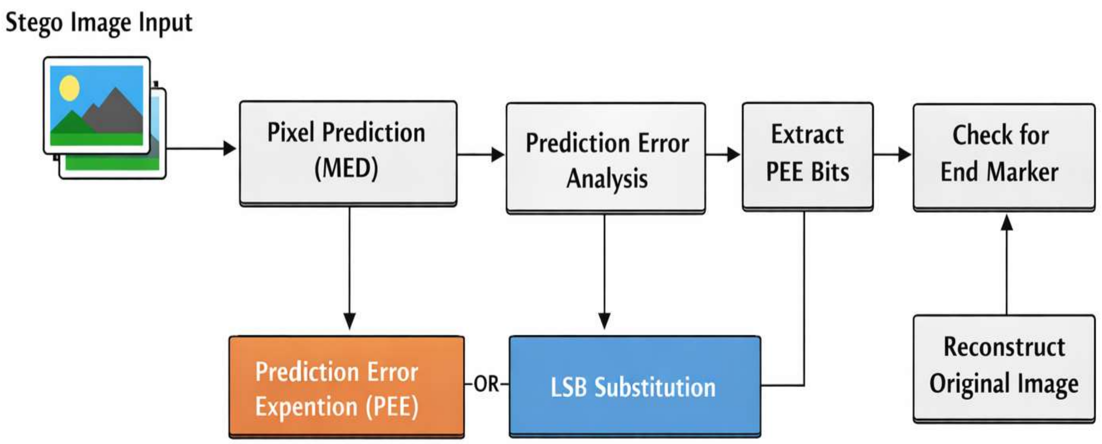

Figure 2: Block Diagram of Extraction Process

Figure 2 represents the extraction phase. The stego image is provided as input. The same MED predictor is applied to maintain consistency. Prediction errors are recalculated to extract embedded bits. Bits embedded via PEE and LSB substitution are sequentially retrieved. Once the termination marker is detected, the secret message is reconstructed. Simultaneously, original pixel values are restored to obtain the exact original image.

Flow Sequence in Figure 2:

Stego Image Input
→ Pixel Prediction
→ Prediction Error Analysis
→ Bit Extraction (PEE + LSB)
→ End Marker Detection
→ Secret Message Recovery
→ Original Image Reconstruction

4. IMPLEMENTATION

The system is implemented using Python programming language with OpenCV and NumPy libraries for efficient pixel manipulation. The implementation consists of two major phases: embedding and extraction.

4.1 Embedding Algorithm

The embedding process begins by reading a lossless PNG cover image. The secret message is converted into ASCII format and then into a binary stream. A termination marker is appended to prevent overflow during extraction.

The image is separated into RGB channels. The MED predictor estimates each pixel value using its upper and left neighboring pixels. Prediction errors are computed for each pixel. If the prediction error lies within the embeddable range, the system embeds one secret bit using Prediction Error Expansion. If not, the least significant bit of the blue channel is modified to embed the secret bit.

This procedure continues sequentially until the entire binary message is embedded. The modified channels are merged to generate the final stego image [14], [6].

4.2 Extraction Algorithm

During extraction, the stego image is processed using the same MED predictor. Prediction errors are recalculated to determine whether PEE embedding was applied. Embedded bits are retrieved accordingly. For LSB-embedded pixels, the least significant bit of the blue channel is extracted.

The system continues extracting bits until the termination marker is identified. The binary sequence is converted back into textual format. Simultaneously, inverse prediction operations restore original pixel values to reconstruct the exact original image [1], [15].

4.3 Flowchart Description

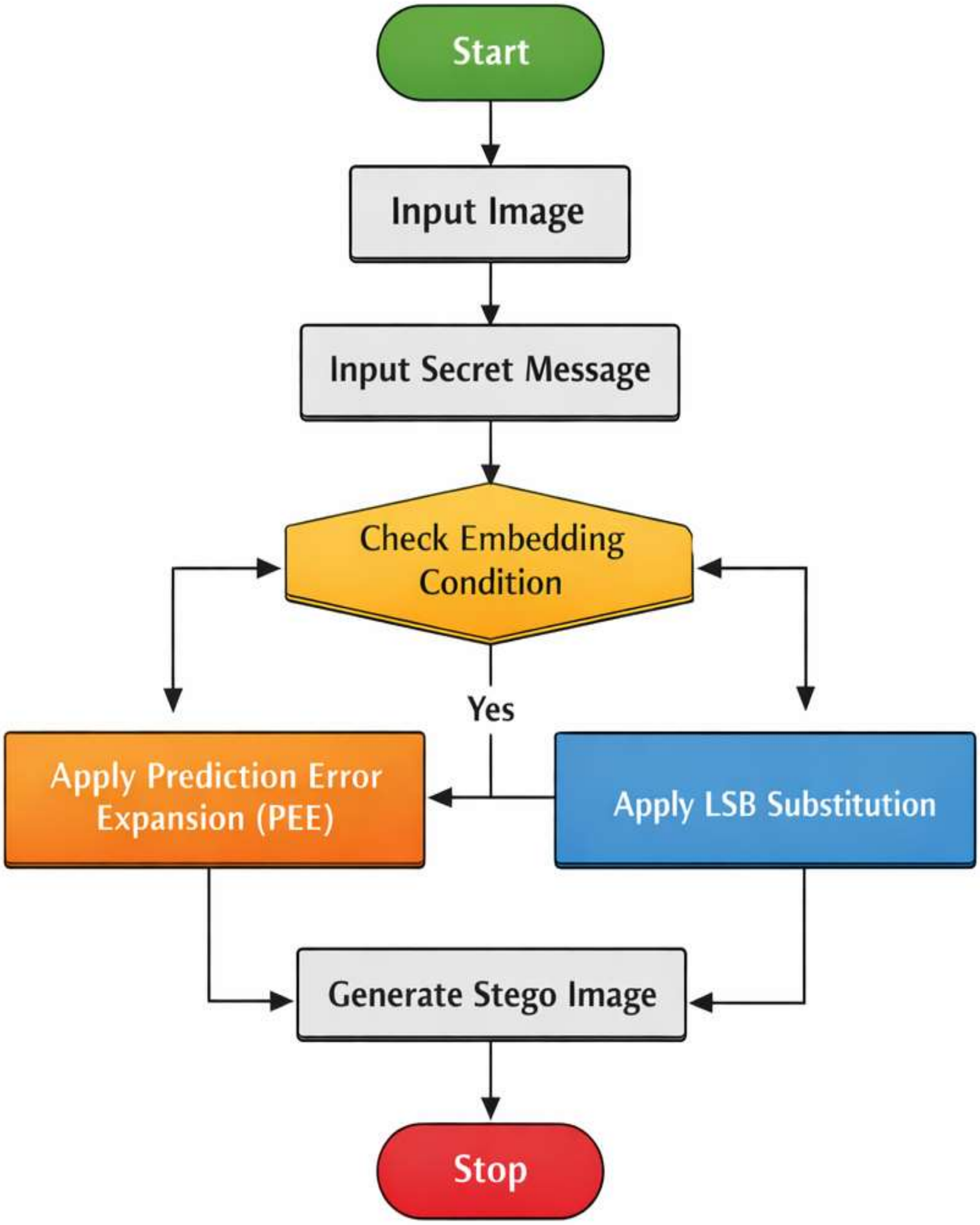

Figure 3: Embedding Flowchart

Figure 3 illustrates the operational workflow of embedding. The process begins with image and message input. The message is converted to binary format, followed by pixel prediction. The system checks embedding conditions and applies either PEE or LSB substitution. The stego image is then generated.

Start
→ Input Image
→ Convert Message to Binary
→ Pixel Prediction
→ Check Embedding Condition
→ Apply PEE or LSB
→ Generate Stego Image
→ Stop

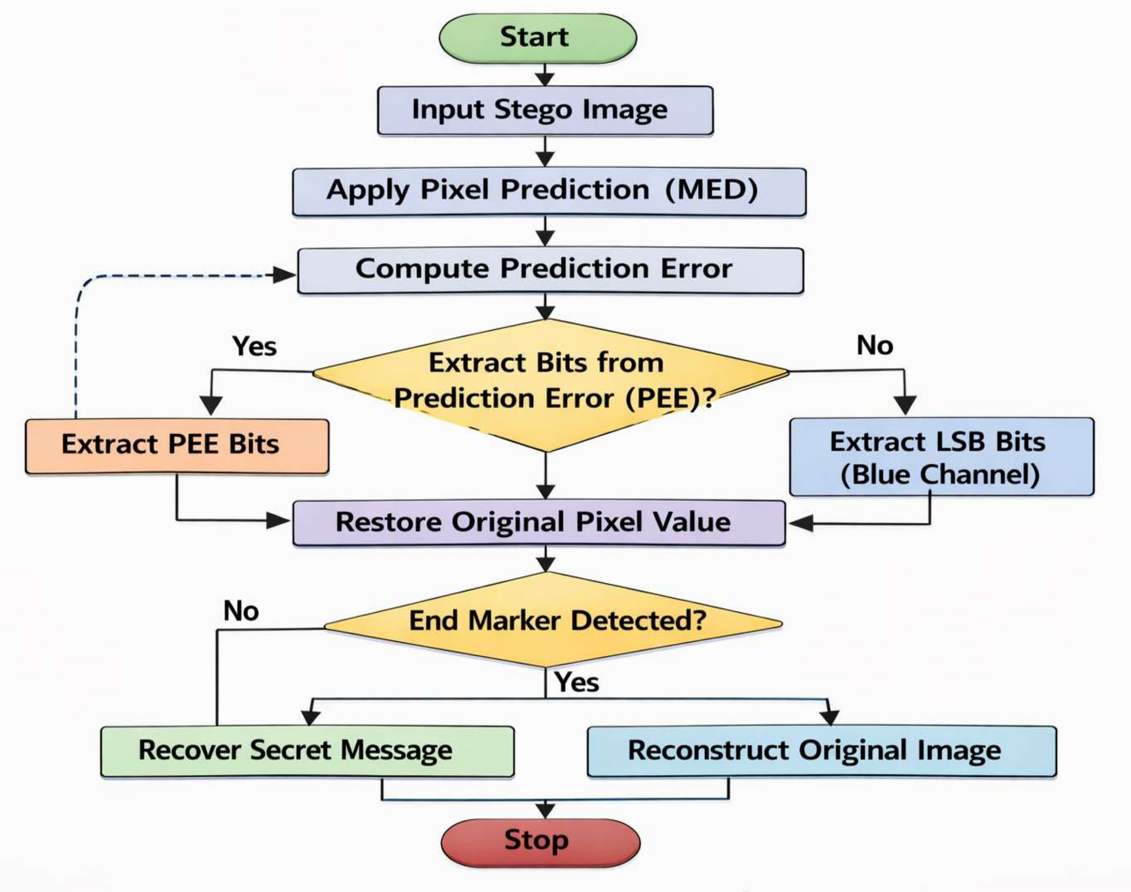

Figure 4: Extraction Flowchart

Figure 4 represents the extraction workflow. The stego image is processed, bits are extracted, termination marker is checked, and both secret message and original image are restored.

Start
→ Input Stego Image
→ Pixel Prediction
→ Extract Bits
→ Check End Marker
→ Restore Pixels
→ Recover Message
→ Stop

5. RESULTS AND DISCUSSION

The performance of the proposed RDH system is evaluated using standard PNG test images. Performance metrics include Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Structural Similarity Index Measure (SSIM) [1], [6].

5.1 Quantitative Results

Table 1: Performance Analysis at Different Payloads

Payload (bits)	PSNR (dB)	MSE	SSIM
10,000	61.8	0.42	0.998
20,000	59.4	0.74	0.996
40,000	57.2	1.10	0.993

The table demonstrates that PSNR values remain above 57 dB even at high payloads. The SSIM values remain close to 1, confirming strong structural similarity between original and stego images.

5.2 Graphical Analysis

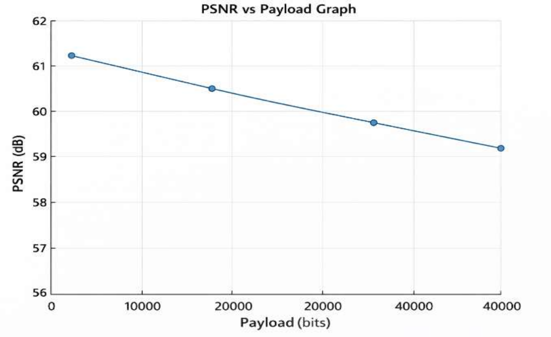

Figure 5: PSNR vs Payload Graph

Figure 5 illustrates that PSNR gradually decreases as payload increases. However, the decline remains moderate due to effective distortion control achieved by prediction-based embedding [14].

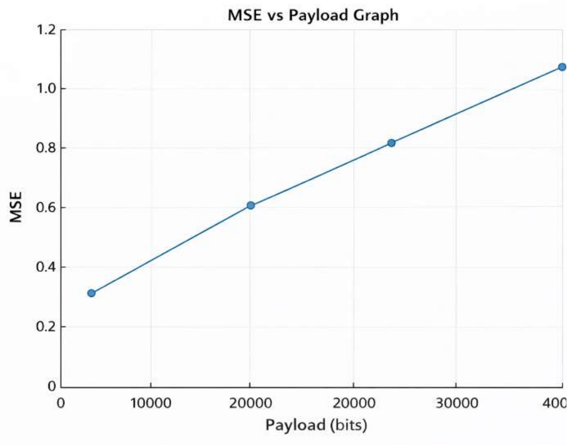

Figure 6: MSE vs Payload Graph

Figure 6 shows a proportional increase in MSE with increasing payload. This aligns with theoretical expectations, as more embedding operations introduce additional pixel modifications.

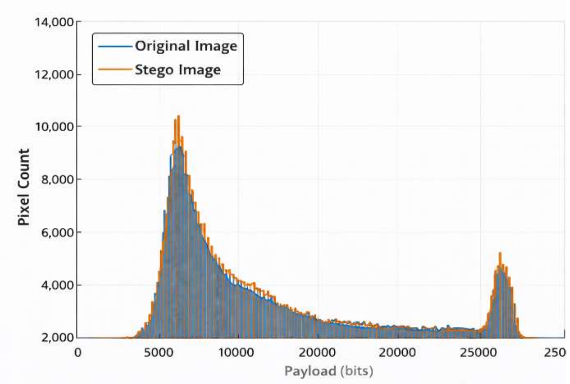

Figure 7: Histogram Comparison of Original and Stego Image

Figure 7 presents histogram comparison between original and stego images. The near-overlapping histograms confirm that statistical properties of the image remain largely unchanged, indicating strong imperceptibility.

5.3 Discussion

The experimental results confirm that the proposed dual embedding framework effectively balances embedding capacity and visual quality. Pure LSB methods provide high capacity but may compromise reversibility and increase detectability [2]. Prediction-based embedding reduces distortion but may limit payload under certain conditions [14].

The integration of PEE and LSB substitution combines the advantages of both techniques. The MED predictor enhances prediction accuracy, resulting in smaller prediction errors and reduced distortion [6]. Comparative observations indicate that the proposed approach achieves performance comparable to advanced prediction-based methods summarized in recent surveys [15].

However, the method is optimized for lossless formats and may not maintain reversibility under lossy compression environments such as JPEG [4].

6. CONCLUSION

This research presented a reversible data hiding method based on dual embedding using Prediction Error Expansion and Least Significant Bit substitution. The system ensures secure embedding while enabling perfect restoration of the original image. Experimental analysis shows PSNR values exceeding 57 dB at high payloads, confirming minimal perceptual distortion. The integration of pixel prediction significantly enhances embedding efficiency compared to standalone LSB or histogram-based methods [1], [15].

7. FUTURE SCOPE

Future research directions include integrating deep learning-based adaptive predictors to further minimize embedding distortion. The system may be extended to encrypted-domain RDH for secure cloud storage applications. Adaptive complexity-based embedding strategies can improve resistance against steganalysis attacks. Additionally, extending the framework to compressed-domain JPEG images would enhance practical applicability [4], [15].

REFERENCES

[1] R. Kumar et al., “Review of Different Prediction Methods for Reversible Data Hiding,” 2023.
[2] K.-H. Jung et al., “Data Hiding Method Using Image Interpolation,” 2009.
[3] H.-W. Tseng et al., “Prediction-Based Reversible Data Hiding,” 2009.
[4] D. Hou et al., “Reversible Data Hiding in JPEG Image Based on DCT Frequency and Block Selection,” 2018.
[5] W. Hong et al., “Reversible Data Hiding for High Quality Images Using Modification of Prediction Errors.”
[6] D. He and Z. Cai, “Reversible Data Hiding for Color Images Using Channel Reference Mapping and Adaptive Pixel Prediction.”
[7] W.-L. Tai et al., “Reversible Data Hiding Based on Histogram Modification of Pixel Differences.”
[11] X. Qu and H. J. Kim, “Pixel-Based Pixel Value Ordering Predictor for High-Fidelity Reversible Data Hiding.”
[14] D. M. Thodi and J. J. Rodríguez, “Prediction Error Expansion.”
[15] X. Zhang et al., “Reversible Data Hiding: Advances in the Past Two Decades.”