An Improved Method for Repeatable Data Hiding

Yipeng Liu and Zichi Wang*

School of Communication and Information Engineering, Shanghai University, Shanghai, 200444, China
Email: liuyipeng@shu.edu.cn (Y.L.); wangzichi@shu.edu.cn (Z.W.)
*Corresponding author

***Abstract*—Repeatable data hiding enables multiple embedding operations on digital images without accumulating distortion. However, existing frameworks reset all the Least Significant Bits (LSBs) to zero for embedding cost assignment. This discards LSB information, reducing imperceptibility. This paper proposes an improved method that achieves distortion invariance without full LSB-zeroing. Instead, it selectively resets only a portion of LSBs, generating a partial LSB-zeroed image. This preserves the original LSB values in non-modifiable regions. By ranking pixels based on embedding costs derived from the partial LSB-zeroed image, our method ensures the same set of pixels is consistently modified across multiple embeddings. As a result, both repeatability and improved imperceptibility are achieved.**

***Keywords*—data hiding, repeatable, imperceptibility**

I. Introduction

Data hiding constitutes a critical aspect within the realm of multimedia security by embedding information imperceptibly into covers, e.g., video, image, audio. Content-adaptive steganography, such as High-pass, Low-pass, and Low-pass (HILL) [1], Wavelet Obtained Weights WOW (WOW) [2], Spatial Universal Wavelet Relative Distortion (SUNIWARD) [3], assigns embedding costs to favor texturally complex regions, improving concealment. Recent advances in deep learning have significantly enhanced data hiding performance, enabling higher embedding capacities and improved security over traditional methods [4–6]. Meanwhile, steganalysis has also evolved, with deep neural networks now offering superior detection accuracy through automated feature extraction [7–8].

Nowadays, perceptual quality metrics are used to assess visual distortion introduced by data hiding techniques [9]. Additionally, multi-platform steganographic systems, such as those employing random-LSB strategies, have been applied to our daily lives [10].

Reversible data hiding techniques allow exact recovery of the original cover image after data extraction [11–16], However, these methods generally require full restoration of the cover before any new embedding, limiting their applicability in scenarios demanding repeatable embedding.

To address this limitation, Wang *et al.* [17] introduced a novel framework that allows sequential embedding operations without requiring prior knowledge of previous modifications. Their approach resets all LSBs of the image to zero and calculates embedding costs based on this LSB-zeroed image. While effective for ensuring repeatability, this approach discards original LSB information, which is essential for enhancing imperceptibility and evading steganalysis. Moreover, embedding costs derived from the LSB-zeroed image are slightly distorted and fail to accurately reflect the statistical characteristics of the original cover.

This paper proposes an improved repeatable method that preserves LSB information in non-modifiable regions while still maintaining repeatability. By partially resetting LSBs and optimizing cost allocation, our method enhances imperceptibility without sacrificing reusability.

The remainder of this paper is organized as follows. Section II details the proposed method. Section III presents experimental results and analysis. Section IV concludes the paper.

II. Improved Repeatable Data Hiding Method

A. Repeatable Data Hiding in [17]

The repeatable data hiding proposed in [17] allows embedding data into the LSBs of digital images while maintaining consistent distortion over multiple embedding iterations. The key idea is to ensure invariant embedding cost and pixel modification locations across embeddings.

Let $\mathbf{I}$ be an $M \times N$ cover image, where each pixel $I(u, v)$, $u \in \{1, \ldots, M\}$, $v \in \{1, \ldots, N\}$) is an 8-bit integer in the range [0, 255]. Denote $\mathbf{I}' = \{I'(u, v)\}$ as the binary LSB plane of $\mathbf{I}$, where $I'(u, v) = I(u, v) \bmod 2$. The framework embeds a capacity of m bits into $\mathbf{I}$, with payload $\gamma = \frac{m}{MN}$ measured in bits per pixel (bpp), which satisfies $0 \le \gamma \le 1$.

Wang *et al.* [17] first deduce that for distortion invariance, the LSB flipping probability $p(u, v)$ of each pixel must either be 0.5 or 0.

Then, they propose an embedding framework as shown in Fig. 1. The initial task is resetting all LSBs of the cover image to zero, producing an LSB-zeroed image. Next, steganographic methods (e.g., HILL, SUNIWARD) compute the embedding cost $\alpha(u, v)$ of each pixel in the LSB-zeroed image. After that, pixels are sorted in

Manuscript received May 20, 2025; revised June 19, 2025; accepted July 23, 2025; published November 25, 2025.

ascending order of $\alpha_0(u, v)$, and the top m pixels are marked as modifiable. Finally, Eq. (1) assigns the final embedding costs $\alpha(u, v)$,

$$\alpha(u,v) = \begin{cases} 1, & I(u,v) \text{ is modifiable} \\ \infty, & \text{otherwise} \end{cases} \quad (1)$$

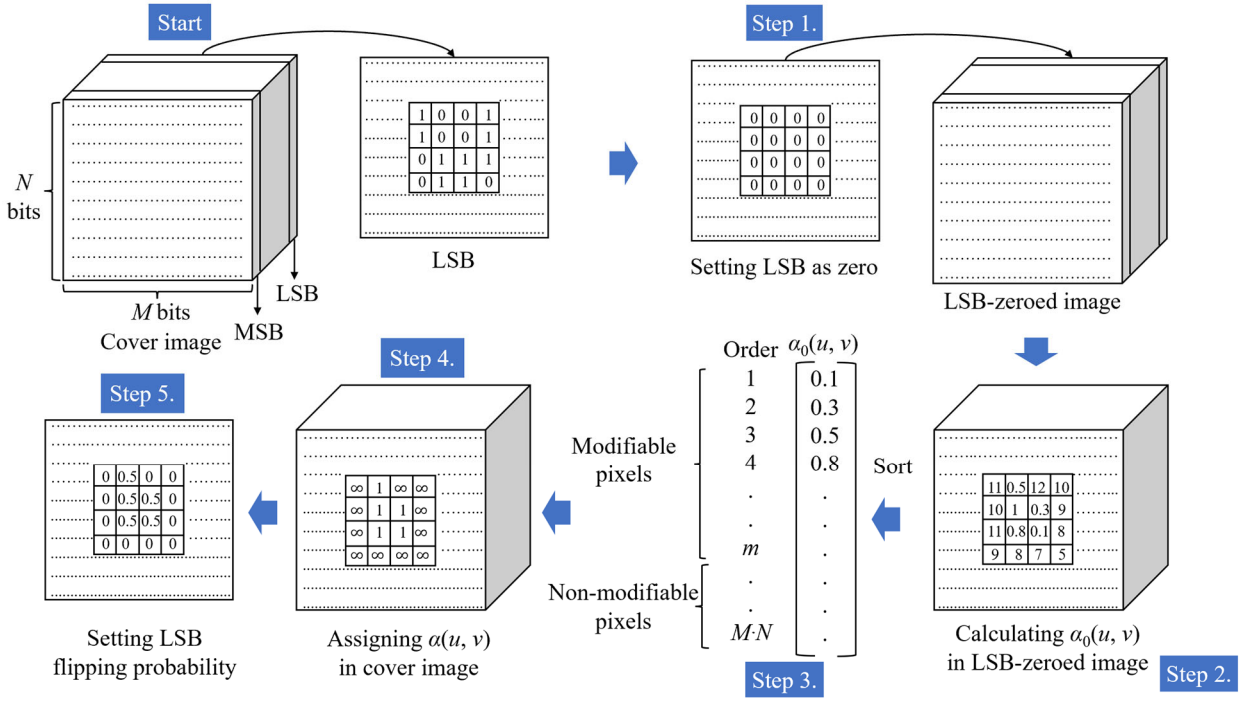


Fig. 1. Embedding framework in [17].

And the additional data is embedded into $\mathbf{I}$ with LSB flipping probability that satisfies Eq. (2),

$$p(u,v) = \begin{cases} 0.5, & I(u,v) \text{ is modifiable} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Although this framework ensures repeatability by fixing the modifiable pixels, it suffers from a major drawback: the complete LSB-zeroing step erases original LSB information. This significantly degrades imperceptibility and introduces distortion into the cost map, which may not accurately reflect the original image statistics.

B. *Improved Method*

To overcome the limitations of [17], LSB information should be preserved under the condition that the location of the modifiable pixels still remains the same during multiple embeddings. That means when the pixels are sorted in ascending order of their embedding cost $\alpha_0(u, v)$ before assigning the final embedding cost $\alpha(u, v)$, the same set of pixels is ranked in the first m positions. This can guarantee the locations of the modifiable pixels are the same during multiple embeddings.

To achieve this, we propose an improved embedding framework that adopts an adaptive partial LSB resetting strategy, as illustrated in Fig. 2. The framework proceeds as follows:

Firstly, the LSB of cover image is set as zero, which generates an LSB-zeroed image. This LSB-zeroed image is utilized to generate an initial embedding cost $\alpha_0(u, v)$ using specific steganographic methods. In this paper, we choose HILL and SUNIWARD as an example to calculate $\alpha_0(u, v)$.

Secondly, all pixels of the cover image are sorted in ascending order of $\alpha_0(u, v)$ and divided into three categories. The top m pixels with minimal $\alpha_0(u, v)$ are marked as modifiable pixels. The next $m \cdot \beta$ pixels are designated as compensating pixels, where β is a compensating parameter. Remaining pixels are non-modifiable pixels.

Thirdly, a partial LSB-zeroed image $\tilde{\mathbf{I}} = \{\tilde{I}(u, v)\}$ is generated by selectively resetting LSBs of modifiable and compensating pixels. Non-modifiable pixels retain their original LSB values. Formally, the partial LSB-zeroed image satisfies.

$$\tilde{I}(u,v) = \begin{cases} I(u,v) - I'(u,v), \\ \text{if } I(u,v) \in \{\text{modifiable} \cup \text{compensating}\} \\ I(u,v), \text{ otherwise} \end{cases} \quad (3)$$

Finally, a refined embedding cost $\alpha_1(u, v)$ is computed from $\tilde{\mathbf{I}}$. The final m modifiable pixels are selected by sorting $\alpha_1(u, v)$. Then, the final embedding costs $\alpha(u, v)$, which ultimately indicate which pixels can be modified, are assigned using Eq. (1). The additional data can be embedded into $\mathbf{I}$ with LSB flipping probability that satisfies Eq. (2).

The use of compensating pixels creates a buffer that mitigates the influence of prior LSB changes in non-modifiable pixels across multiple embeddings. But this will also lead to the distortion of embedding cost in turn. Thus, an equilibrium point between imperceptibility and repeatability should be found. More details about the compensating parameter will be discussed in Section III.

The process of adding additional data into cover image begins with the data embedder generating a matrix $\mathbf{H} \in \{0, 1\}^{m \times MN}$. The matrix $\mathbf{H}$ is constructed to have full rank, i.e., $R(\mathbf{H}) = m$, ensuring the system of linear equations defined by $\mathbf{H}$ is solvable. The LSB of stego image is systematically vectorized into a binary sequence.

$$\mathbf{I}' = [I'(1), I'(2), \ldots, I'(MN)]^T \in \{0, 1\}^{MN \times 1} \quad (4)$$

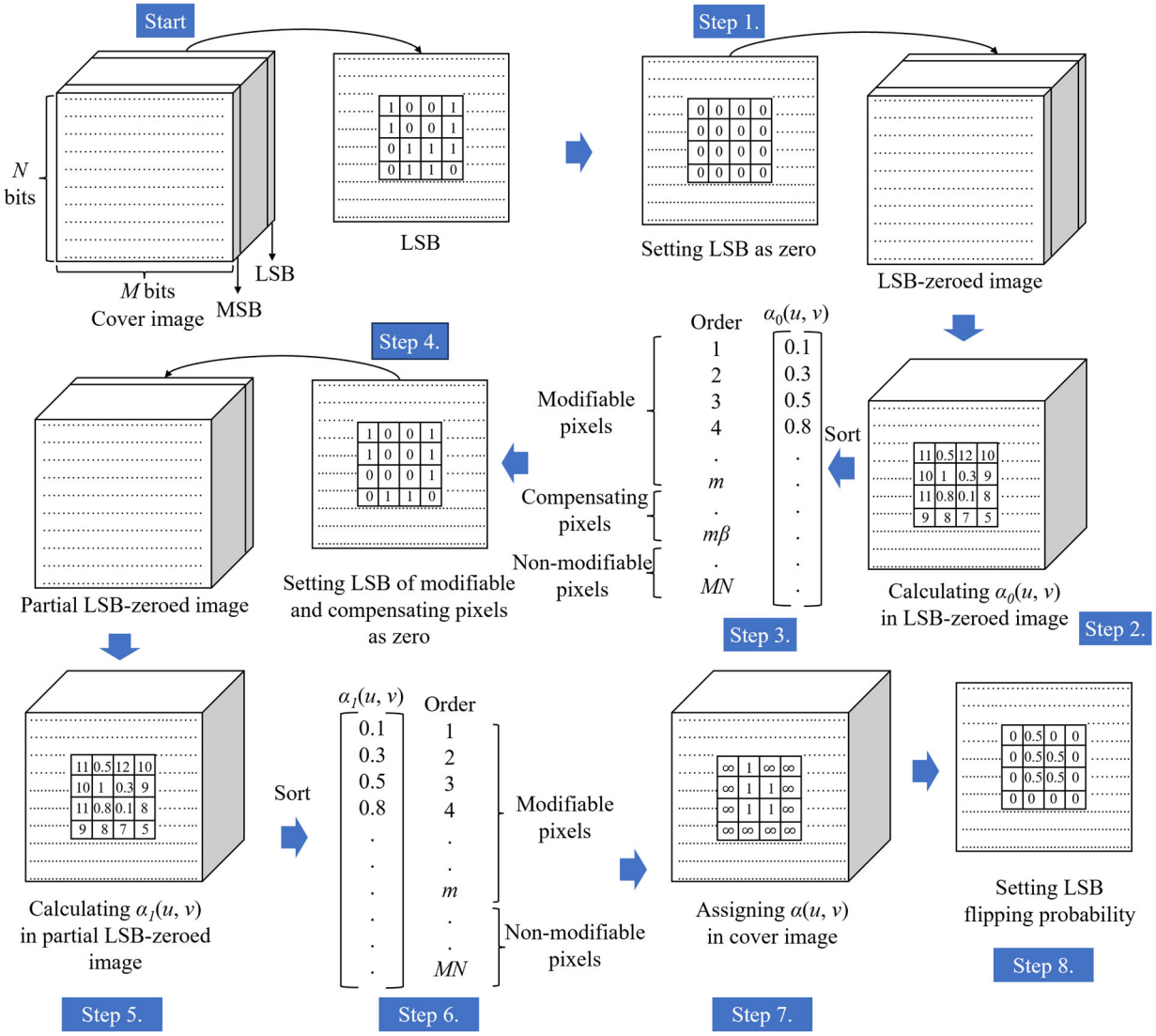


Fig. 2. Improved embedding framework.

where each element corresponds to the LSB of a pixel at position $\left(\left\lfloor\frac{r-1}{N}\right\rfloor + 1, r - \left\lfloor\frac{r-1}{N}\right\rfloor N\right)$, where $r \in \{1, 2, \ldots, MN\}$ and "$\lfloor\cdot\rfloor$" stands for the floor rounding operation. During embedding, the secret messages $\boldsymbol{w} = [w(1), w(2), \ldots, w(m)]^T$ can be embedded into $\mathbf{I}$ by enforcing the Eq. (5).

$$\mathbf{w} = \mathbf{H}\mathbf{I}' \quad (5)$$

To satisfy this equation, each row of $\mathbf{H}$ is treated as a parity-check equation, the linear combination should be satisfied.

$$w(t) = \sum_{r=1}^{MN} h(t, r) \cdot I'(r) \quad (6)$$

where $t \in \{1, 2, \ldots, m\}$.

If this condition is not met, the algorithm flips the LSB of a modifiable pixel $I(r)$ where $h(t, r) = 1$, iterating until all equations are satisfied. As mentioned above, the modifiable pixels are predefined based on their embedding costs.

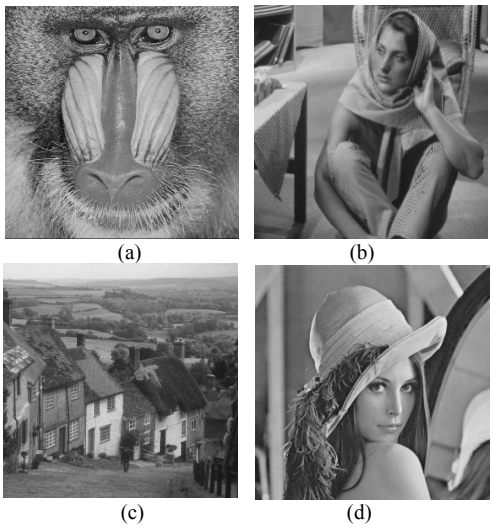
For extraction, the data embedder sends a data hiding key to the receiver, from which the same $\mathbf{H}$ can be reconstructed. The additional data $\mathbf{w}$ can be directly computed through Eq. (5) from the stego image's LSB vector.

III. Experimental Result

To validate the improvement of our method, we conduct comprehensive experiments under standardized conditions. Initially, the experimental environments are set up. Subsequently, we ascertain the values of compensating parameter β. Finally, we analyze the quality of stego image and make comparison with the method in Ref. [17].

A. Experiment Setup

We conducted experiments using the UCID dataset [18]. This dataset contains 1338 uncompressed color images, each sized 512×384 pixels. All images were converted to grayscale to serve as cover images. Additionally, several standard test images widely used in image processing were included, as shown in Fig. 3. For our proposed method, embedding costs were assigned following the procedure in Subsection II.B. Specifically, the initial costs were calculated using the HILL and SUNIWARD. For comparative evaluation, we employed the repeatable data hiding method described in [17].



(a) (b) (c) (d)

Fig. 3. Several popular images used in image processing (a) Baboon; (b) Barbara; (c) Goldhill; (d) Lena.

B. Optimization of Compensating Parameter

To balance the trade-off between repeatability and imperceptibility, we systematically determine the optimal compensating parameter β through a data-driven approach. According to our following experiment results, the compensating parameter β should be dynamically adjusted across payloads to maintain optimal imperceptibility while preserving repeatability.

As defined in Subsection II.B, β governs the proportion of compensating pixels reset during partial LSB-zeroing. Let $m = \gamma \cdot MN$ denote the number of modifiable pixels. The number of compensating pixels is $m \cdot \beta = \beta \cdot \gamma \cdot MN$. To ensure feasibility, the combined count of modifiable and compensating pixels must satisfy.

$$\gamma \cdot MN + \beta \cdot \gamma \cdot MN < MN \tag{7}$$

Then we have

$$\beta < \frac{1}{\gamma} - 1 \tag{8}$$

Since $\beta \geqslant 0$, the valid range is $0 \leqslant \beta < \frac{1}{\gamma} - 1$. The valid ranges of β for selected payload values are listed in Table I. JSD (Jensen-Shannon Divergence) [19] is adopted to quantify the statistical deviation between cover and stego images. The range of JSD is $0 \leq \text{JSD} \leq 1$. Lower JSD means less difference between the cover image and stego image. Thus, our work is to determine a proper β to make JSD of the stego image as low as possible.

In this experiment, the payload for each image is set as $\gamma \in \{0.1, 0.2, \ldots, 1\}$ bpp. For each payload γ, β is evaluated across its valid range with a granularity of 0.01. For each (γ, β) pair, we execute sequentially $K = 5$ times embeddings on $J = 100$ UCID images. After the k-th embedding ($k = 1, 2, \ldots, 5$) on the j-th image ($j = 1, 2, \ldots, 100$), the JSD of the stego image $\text{JSD}_{j,k}(\gamma, \beta)$ is computed. For each image j and embedding time k with certain payload γ, identify β that minimizes JSD as:

$$\beta_{j,k}(\gamma) = \min\{\text{JSD}_{j,k}(\gamma, \beta)\} \tag{9}$$

After $\beta_{j,k}(\gamma)$ across 100 images and 5 embedding times are collected, the final $\hat{\beta}(\gamma)$ is calculated as:

$$\hat{\beta}(\gamma) = \frac{1}{J \cdot K} \sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{j,k}(\gamma) \tag{10}$$

The result of $\hat{\beta}(\gamma)$ is shown in Table I. However, in real-world scenarios, the payload is a various value tailored to the size of additional information. Thus, to generalize the empirically determined optimal compensating parameter from discrete payloads $\gamma \in \{0.1, 0.2, \ldots, 1.0\}$ bpp to a continuous domain $\gamma \in (0, 1]$, piecewise linear interpolation is employed. This method constructs a continuous function $\beta(\gamma)$ by linearly connecting adjacent discrete data points $(\gamma_i, \hat{\beta}_i)$. Mathematically, for any $\gamma \in [\gamma_i, \gamma_{i+1}]$, the interpolated value is defined as:

$$\beta(\gamma) = \hat{\beta}_i + \frac{\hat{\beta}_{i+1} - \hat{\beta}_i}{\gamma_{i+1} - \gamma_i}(\gamma - \gamma_i) \tag{11}$$

where γ_i and γ_{i+1} are consecutive payload values from Table I.

TABLE I. VALID RANGES AND OPTIMAL β FOR DIFFERENT PAYLOADS

Payload (bpp)	Range of β	$\hat{\beta}(\gamma)$
0.1	[0, 9.00)	3.48
0.2	[0, 4.00)	1.77
0.3	[0, 2.33)	1.25
0.4	[0, 1.50)	0.73
0.5	[0, 1.00)	0.56
0.6	[0, 0.67)	0.34
0.7	[0, 0.43)	0.22
0.8	[0, 0.25)	0.13
0.9	[0, 0.11)	0.06
1.0	0	0

As shown in Fig. 4, the curve of $\beta(\gamma)$ strictly respects the experimentally observed monotonic relationship between β and γ, where β decreases as γ increases, reflecting the need to minimize compensatory resets at higher payloads to avoid statistical anomalies. Linear interpolation maintains this trend by ensuring negative slopes between adjacent points, consistent with empirical results. Additionally, piecewise linear interpolation avoids overfitting noise and unnecessary oscillations.

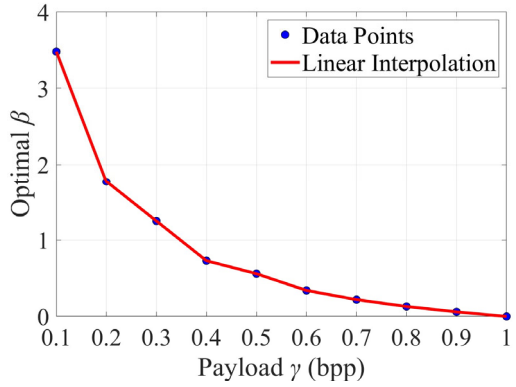


Fig. 4. The curve of $\beta(\gamma)$.

C. Imperceptibility

The imperceptibility of stego images is rigorously evaluated using five complementary metrics: PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity Index Measurement) [20], AB-SSIM (Attention-Based SSIM), KLD (Kullback-Leibler Divergence), and JSD. The ranges of PSNR, SSIM, AB-SSIM, and KLD are $0 < \text{PSNR} \leq +\infty$, $0 \leq \text{SSIM} \leq 1$, $0 \leq \text{AB-SSIM} \leq 1$, and $0 < \text{KLD} \leq +\infty$, respectively. AB-SSIM enhances SSIM by using a spatial attention module, which focuses on 'where' is an informative part [21]. Spatial attention assigns higher weights to perceptually critical regions. This weighting strategy better aligns with human visual perception by emphasizing distortions in visually salient areas. Higher PSNR, SSIM, and AB-SSIM indicate better preservation of structural and luminance information, while lower JSD and KLD reflect minimized statistical deviations between cover and stego images.

Experimental results, as detailed in Tables II and III, demonstrate the effectiveness of our method when using the HILL cost assignment method. Across payloads ranging from 0.1 to 1.0 bpp, our method consistently achieves PSNR values exceeding 50 dB and SSIM values approaching the ideal value of 1.0, with UCID dataset averages maintaining this performance trend.

The trade-off between imperceptibility and embedding capacity is observed in Tables II and III. Higher payloads necessitate more modifications, leading to gradual reductions in PSNR and SSIM. Our method maintains practical usability even at maximum capacity (1.0 bpp).

TABLE II. PSNR (dB) OF OUR METHOD WITH PAYLOADS {0.1, 0.2, …, 1} BPP

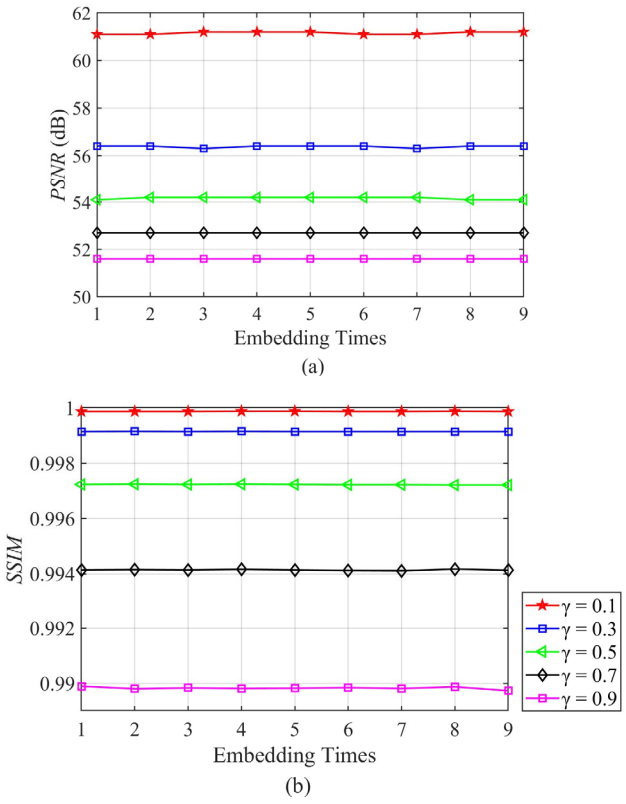
Payload (bpp)	Baboon	Barbara	Goldhill	Lena	UCID
0.1	61.20	61.10	61.20	61.10	61.14
0.2	58.10	58.10	58.20	58.10	58.12
0.3	56.40	56.40	56.40	56.40	56.39
0.4	55.20	55.10	55.10	55.10	55.10
0.5	54.20	54.20	54.10	54.20	54.15
0.6	53.40	53.30	53.40	53.40	53.38
0.7	52.70	52.70	52.70	52.70	52.70
0.8	52.10	52.10	52.10	52.10	52.10
0.9	51.60	51.60	51.60	51.60	51.60
1.0	51.20	51.20	51.20	51.20	51.20

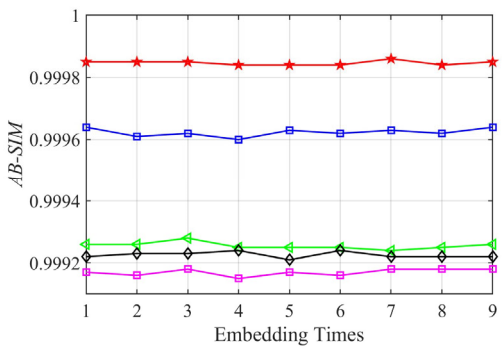
TABLE III. SSIM OF OUR METHOD WITH PAYLOADS {0.1, 0.2, …, 1} BPP

Payload (bpp)	Baboon	Barbara	Goldhill	Lena	UCID
0.1	0.9999	0.9999	0.9999	0.9999	0.9999
0.2	0.9999	0.9999	0.9997	0.9995	0.9995
0.3	0.9999	0.9998	0.9996	0.9991	0.9990
0.4	0.9999	0.9994	0.9994	0.9984	0.9980
0.5	0.9998	0.9983	0.9990	0.9972	0.9964
0.6	0.9997	0.9971	0.9987	0.9958	0.9937
0.7	0.9996	0.9959	0.9982	0.9941	0.9888
0.8	0.9994	0.9947	0.9978	0.9921	0.9809
0.9	0.9990	0.9931	0.9971	0.9899	0.9692
1.0	0.9985	0.9912	0.9934	0.9877	0.9509

In our proposed method, by setting proper compensating parameter β historical modifications are partially erased. This partial erasure offsets distortion arising from varying embedding costs across operations. Consequently, overall distortion from data hiding is minimized. As a result, the PSNR, SSIM, and AB-SSIM values will hardly decrease during multiple embeddings. To validate this, we measure these metrics on Lena with payloads $\gamma \in \{0.1, 0.3, \ldots, 0.9\}$ bpp under sequential embeddings. As demonstrated in Fig. 5, all metrics remain essentially invariant during 9 embeddings, confirming the method's repeatability. The slight fluctuations are attributable to the inherent randomness of embedded binary sequences.

Furthermore, comparative analysis of Fig. 5(b) and 5(c) reveals that AB-SSIM maintains higher values than SSIM, especially at larger payloads, demonstrating better imperceptibility to visually salient regions. This is because our method restricts modifications to low-risk areas, which preserves structural integrity in visually salient regions weighted heavily by AB-SSIM.

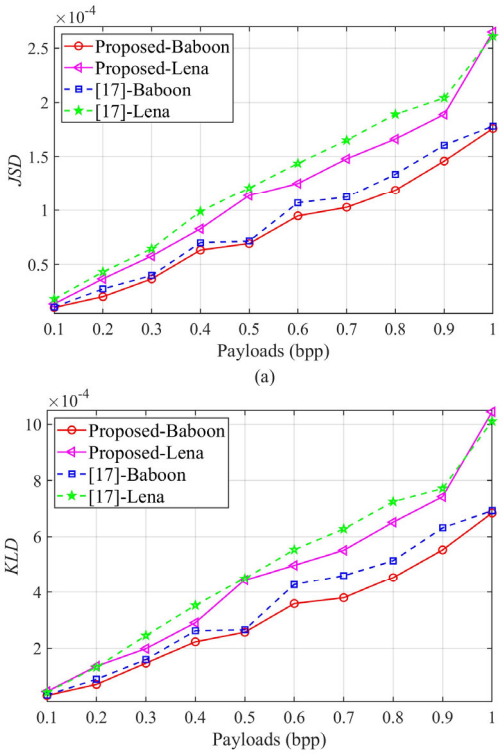




(c)

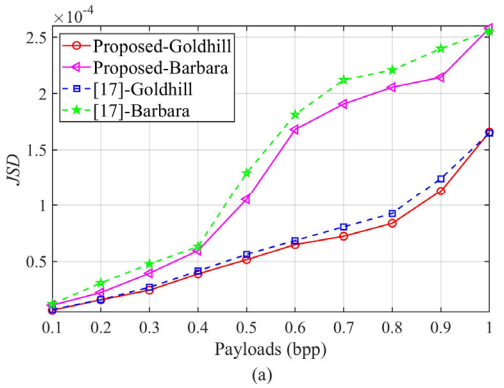
Fig. 5. Imperceptibility of our method for (a) PSNR; (b) SSIM; (c) AB-SSIM on Lena with multiple embeddings.

To evaluate the imperceptibility improvements of our method, we also provide a group of imperceptibility comparisons with the method in [17], shown in Figs. 6–9 and Tables IV and V.

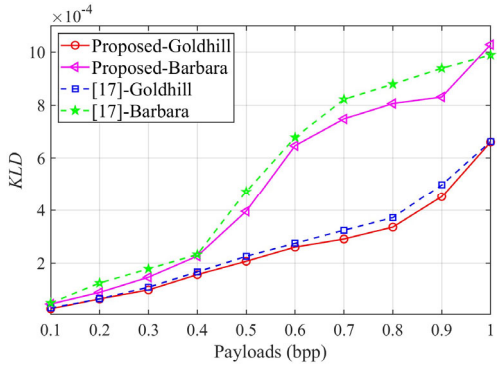


(a) (b)

Fig. 6. Imperceptibility comparisons of our method and the method in [17] for (a) JSD; (b) KLD tested on Baboon, Lena.

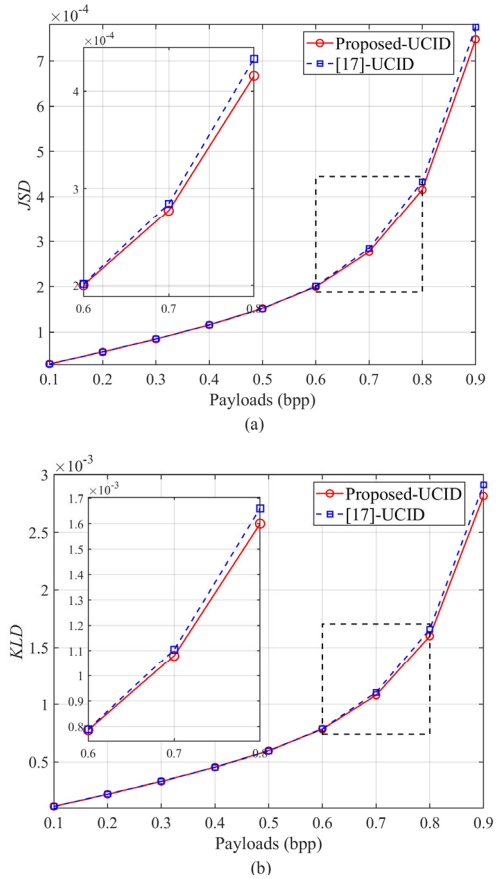


(a)



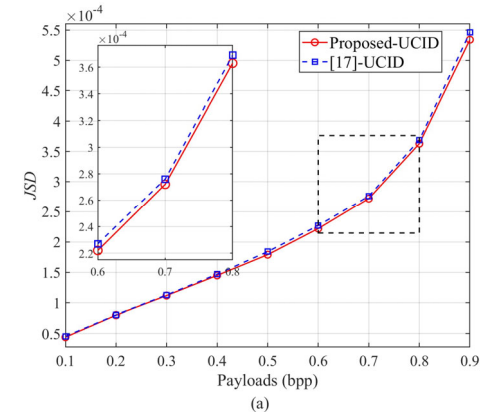
(b)

Fig. 7. Imperceptibility comparisons of our method and the method in [17] for (a) JSD; (b) KLD tested on Goldhill, Barbara.



(a) (b)

Fig. 8. Imperceptibility comparisons of our method and the method in [17] for average (a) JSD; (b) KLD tested on the image in UCID using HILL.



(a)

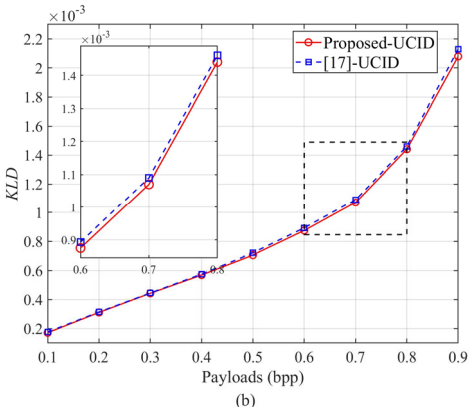


Fig. 9. Imperceptibility comparisons of our method and the method in [17] for average (a) JSD; (b) KLD tested on the image in UCID using SUNIWARD.

The line plots of Figs. 8 and 9 exhibited significant overlap due to the excessively large scale of the data points. Thus, we also employ a data table (shown in Tables IV and V) to present the numerical results.

It can be seen that our method achieves significantly lower JSD and KLD values across all tested payloads $\gamma \in \{0.1, 0.2, \ldots, 1\}$ bpp. The reason is that the adaptive partial LSB resetting strategy, which preserves the information of the original LSB plane in non-modifiable regions, unlike the global LSB-zeroing in [17]. Therefore, the imperceptibility of our method is better than the method in [17].

To validate generalizability beyond HILL, we replicated the imperceptibility analysis using the SUNIWARD cost assignment. Fig. 9. presents the average JSD and KLD results for SUNIWARD on the UCID dataset. Consistent with the HILL results in Fig. 8, our method achieves lower JSD and KLD values compared to [17] across all payloads. Numerical results are detailed in Tables IV and V. This demonstrates that the proposed partial LSB resetting strategy effectively preserves statistical characteristics regardless of the underlying cost function

To statistically validate the imperceptibility improvements, we conducted Wilcoxon signed-rank tests on JSD and KLD metrics at payload $\gamma = 0.5$ bpp using HILL cost assignment. The tests were performed on UCID images comparing our method against [17]. The null hypothesis (H_0) stated no difference in distortion between two methods, while the alternative hypothesis (H_1) asserted lower distortion with our method.

As shown in Table VI, significantly negative z-values with $p < 0.001$ confirm our method produces substantially lower distortion. Large effect sizes ($|r| > 0.8$) and negative median differences demonstrate not only statistical significance but also practical relevance of these improvements.

TABLE IV. AVERAGE JSD TESTED ON THE IMAGE IN UCID

Steganographic Methods	Payload (bpp)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
HILL	[17]-UCID	2.978×10^{-5}	5.670×10^{-5}	8.488×10^{-5}	1.160×10^{-4}	1.523×10^{-4}	2.013×10^{-4}	2.844×10^{-4}	4.333×10^{-4}	7.747×10^{-4}
	Proposed-UCID	2.969×10^{-5}	5.654×10^{-5}	8.471×10^{-5}	1.158×10^{-4}	1.517×10^{-4}	2.001×10^{-4}	2.771×10^{-4}	4.157×10^{-4}	7.473×10^{-4}
SUNIWARD	[17]-UCID	4.551×10^{-5}	8.049×10^{-5}	1.131×10^{-4}	1.467×10^{-4}	1.832×10^{-4}	2.265×10^{-4}	2.748×10^{-4}	3.677×10^{-4}	5.455×10^{-4}
	Proposed-UCID	4.420×10^{-5}	8.017×10^{-5}	1.129×10^{-4}	1.451×10^{-4}	1.793×10^{-4}	2.224×10^{-4}	2.730×10^{-4}	3.672×10^{-4}	5.344×10^{-4}

TABLE V. AVERAGE KLD TESTED ON THE IMAGE IN UCID

Steganographic Methods	Payload (bpp)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
HILL	[17]-UCID	1.115×10^{-4}	2.212×10^{-4}	3.331×10^{-4}	4.552×10^{-4}	5.992×10^{-4}	7.891×10^{-4}	1.104×10^{-3}	1.660×10^{-3}	2.915×10^{-3}
	Proposed-UCID	1.114×10^{-4}	2.206×10^{-4}	3.328×10^{-4}	4.544×10^{-4}	5.962×10^{-4}	7.855×10^{-4}	1.080×10^{-3}	1.600×10^{-3}	2.816×10^{-3}
SUNIWARD	[17]-UCID	1.757×10^{-4}	3.138×10^{-4}	4.449×10^{-4}	5.753×10^{-4}	7.207×10^{-4}	8.923×10^{-4}	1.081×10^{-3}	1.442×10^{-3}	2.127×10^{-3}
	Proposed-UCID	1.710×10^{-4}	3.116×10^{-4}	4.435×10^{-4}	5.694×10^{-4}	7.065×10^{-4}	8.758×10^{-4}	1.075×10^{-3}	1.440×10^{-3}	2.086×10^{-3}

TABLE VI. WILCOXON SIGNED-RANK TEST RESULTS

Metric	z-value	*p*-value	r	Median Difference
JSD	−3.734	0.0002	−0.880	-9.13×10^{-6}
KLD	−3.418	0.0006	−0.806	-1.87×10^{-5}

D. Undetectability

We further assess the undetectability of our method against conventional steganalysis using Spatial Rich Model (SRM) [22] and Subtractive Pixel Adjacency Matrix (SPAM) [23] feature sets. For each payload level $\gamma \in \{0.3, 0.5\}$ bpp (commonly adopted in steganographic benchmarking), we embed data using our proposed method and the method in [17] respectively. The average SRM and SPAM feature values across all test images are then computed for both methods.

A portion of the SRM and SPAM feature values is illustrated in Figs. 10 and 11. The near-overlapping points indicate no significant divergence between the two methods. This demonstrates that our method preserves the same level of undetectability as [17], despite eliminating full LSB-zeroing. This is because Wang [17] already achieves high security, the effect of our refinement is relatively limited. Consequently, our method maintains undetectability while addressing the imperceptibility limitations of the original framework.

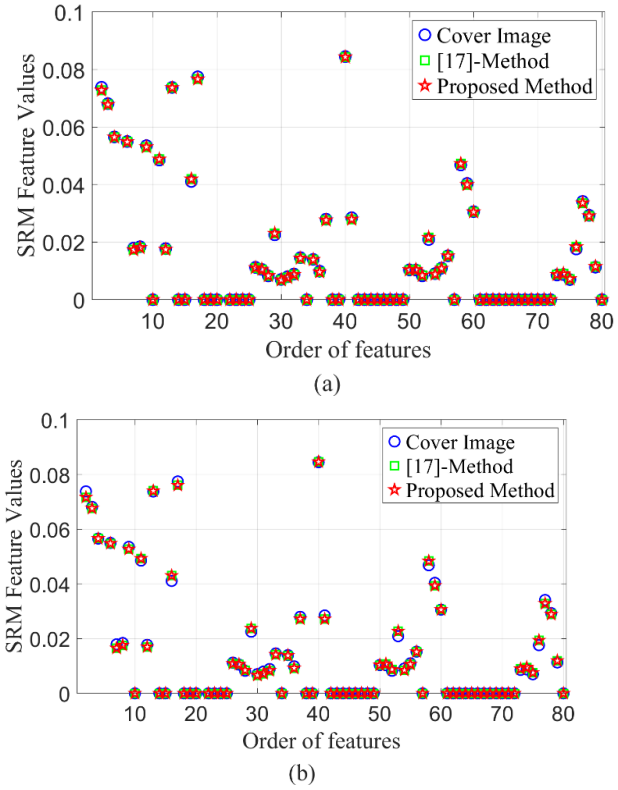


Fig. 10. Comparisons of SRM feature values at payload (a) 0.3 bpp (b) 0.5 bpp.

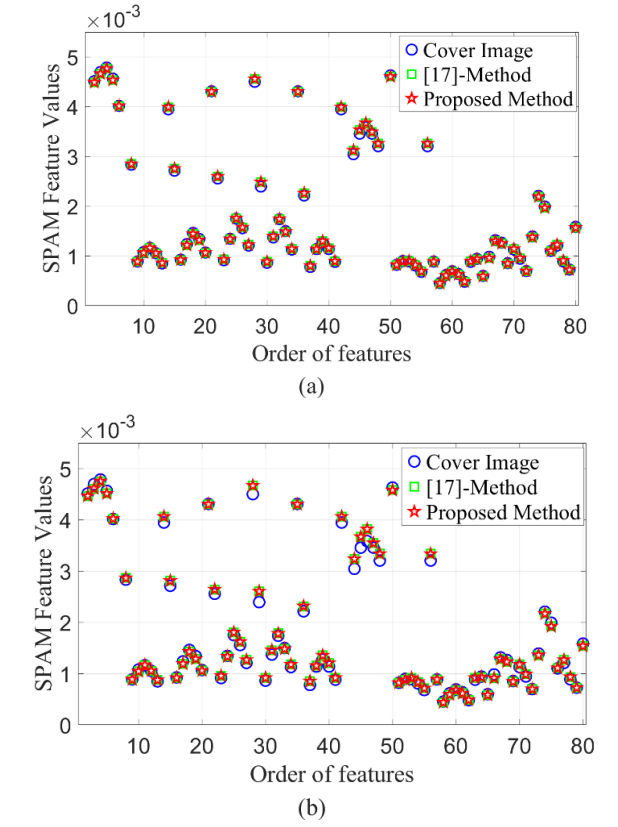


Fig. 11. Comparisons of SPAM feature values at payload (a) 0.3 bpp (b) 0.5 bpp.

IV. Conclusion

This paper presents an improved framework for repeatable data hiding that addresses the imperceptibility limitations of prior methods. By adopting a selective LSB resetting strategy, our approach retains the original LSB values in non-modifiable regions, significantly reducing visual and statistical distortion. The introduction of compensating pixels ensures that embedding cost consistency is maintained across multiple embeddings, thereby preserving repeatability.

In future work, we will explore cross-media repeatable embedding, extending the current method beyond static images to audio, video, and other modalities. Moreover, as shown in Table I, the compensating parameter β currently only depends on the load γ and does not consider the difference in image content. Thus, designing an image-dependent β estimation strategy remains an important research direction. Finally, since our current undetectability evaluation is based on traditional feature sets such as SRM and SPAM, future work will focus on assessing robustness against advanced deep learning-based classifiers.

Repeatable data hiding enhances the flexibility of secure multimedia communication frameworks. However, it also poses ethical concerns. Responsible development and deployment are essential to ensure these technologies serve legitimate privacy and security applications.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Yipeng Liu conducted the research, implemented the proposed method, and performed the experiments. Zichi Wang supervised the project, contributed to the design of the methodology, and provided guidance throughout the research. Yipeng Liu drafted the manuscript, and Zichi Wang revised it critically for important intellectual content. All authors have read and approved the final version of the manuscript.

Funding

This work was supported in part by Natural Science Foundation of China under Grant 62376148.

References

[1] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2014, pp. 4206–4210, Oct. 2014.

[2] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP J. Inf. Secur.*, vol. 2014, no. 1, pp. 1–13, Jan. 2014.

[3] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2012, pp. 234–239.

[4] Z. Guan *et al.*, "DeepMIH: Deep Invertible network for multiple image hiding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 372–390, Jan. 2022.

[5] S.-P. Lu, R. Wang, T. Zhong, and P. L. Rosin, "Large-capacity image steganography based on invertible neural networks," in *Proc.*

2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2021, pp. 10811–10820.
[6] R. Zhang, S. Dong, and J. Liu, "Invisible steganography via generative adversarial networks," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 8559–8575, Dec. 2018.
[7] I. T. Ahmed, B. T. Hammad, and N. Jamil, "Image steganalysis based on pretrained convolutional neural networks," in *Proc. 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA)*, Selangor, Malaysia, 2022, pp. 283–286.
[8] P. Li, Y. Li, H. Wang, and C. Liu, "Research on steganalysis of digital image based on deep learning," in *Proc. 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, Mar. 2021, pp. 528–534.
[9] A. Pramanick, D. Megha, and A. Sur, "Attention-based spatial-frequency information network for underwater single image super-resolution," in *Proc. ICASSP 2024—2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Seoul, Korea, 2024, pp. 3560–3564.
[10] G. A. Tibor and J. Katona, "Development of multi-platform steganographic software based on random-LSB," *Program Comput. Soft,* vol. 49, pp. 922–941, 2023.
[11] M. Lin and S. Xiang, "Efficient CNN prediction with smoothness factor for reversible data hiding," *IEEE Signal Processing Lett.*, pp. 1–5, Jan. 2025.
[12] R. Hu and S. Xiang, "CNN prediction based reversible data hiding," *IEEE Signal Process. Lett.*, vol. 28, pp. 464–468, 2021.
[13] R. Hu and S. Xiang, "Reversible data hiding by using CNN prediction and adaptive embedding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 10196–10208, Dec. 2022.
[14] X. Yang and F. Huang, "New CNN-based predictor for reversible data hiding," *IEEE Signal Process. Lett.*, vol. 29, pp. 2627–2631, Jan. 2022.
[15] M. Wu and S. Xiang, "An efficient CNN-based prediction for reversible data hiding," in *Proc. 5th ACM Int. Conf. Multimedia Asia*, 2023, pp. 1–5.
[16] H. Rao, S. Weng, L. Yu, L. Li, and G. Cao, "High-precision reversible data hiding predictor: UCANet," *IEEE Signal Process. Lett.*, vol. 31, pp. 2155–2159, Jan. 2024.
[17] Z. Wang, G. Feng and X. Zhang, "Repeatable data hiding: Towards the reusability of digital images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 1, pp. 135–146, Jan. 2022.
[18] G. Schaefer and M. Stich, "UCID—An uncompressed colour image database," in *Proc. Conf. Storage Retr. Methods Appl. Multimedia*, San Jose, CA, USA, Jan. 2004, pp. 472–480.
[19] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, Jan. 1991.
[20] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
[21] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. the European Conference on Computer Vision (ECCV)*, 2018, pp. 3–19.
[22] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 868–882, Jun. 2012.
[23] T. Pevny, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 215–224, Jun. 2010.