Secure Sudoku Mapping for Dual-Image Reversible Data Hiding

DOI: https://doi.org/10.54654/isj.v2i28.6419

Cao Thi Luyen

***Abstract*— Reversible data hiding in encrypted images (RDHEI) often creates embedding room by preserving or preprocessing image redundancy, thereby coupling the hiding layer to a specialized encryption model. This paper presents a dual-image method that works directly on AES-CTR ciphertext. Its design rests on four elements: a keyed Sudoku coordinate code for carrying two base-8 digits per accepted pair; exact reconstruction of each ciphertext pair from inter-block displacement; AES-GCM framing for payload confidentiality and integrity; and an adaptive movement threshold that selects the lowest-distortion mappings able to accommodate the complete frame. The data hider requires neither plaintext nor the image-encryption key, and no location map is transmitted. On sixteen 512×512 grayscale images, the method achieves a mean maximum net rate of 1.4187 bits per transmitted pixel. At ER_t=1.0, the two marked outputs reach mean PSNR values of 47.65 dB, and all 160 image-rate trials yield zero payload error and pixel-exact recovery. An independent implementation of the embedding and recovery layer of Venkatesh et al. is also examined; published and reproduced results are reported separately where the original specification leaves ambiguity. The experiments show that the proposed architecture provides predictable payload, low carrier-domain distortion, authenticated framing, and exact recovery while retaining standard image encryption.**

***Tóm tắt*— Giấu tin thuận nghịch trong ảnh mã hóa thường tạo không gian nhúng bằng cách bảo toàn hoặc tiền xử lý dư thừa ảnh, khiến lớp giấu tin phụ thuộc vào một mô hình mã hóa chuyên biệt. Bài báo đề xuất một phương pháp hai ảnh hoạt động trực tiếp trên ảnh mã AES-CTR. Phương pháp được xây dựng trên bốn thành phần chính: mã tọa độ Sudoku sinh từ khóa để mang hai chữ số cơ số 8 trên mỗi cặp hợp lệ; khôi phục chính xác từng cặp byte mã từ độ dịch chuyển giữa các khối; khung tải tin được bảo mật và xác thực bằng AES-GCM; và ngưỡng dịch chuyển thích nghi để lựa chọn các phép ánh xạ có độ méo thấp nhất nhưng vẫn đủ chứa toàn bộ khung dữ liệu. Bên giấu tin không cần biết ảnh rõ hoặc khóa mã hóa ảnh, đồng thời phương pháp không truyền bản đồ vị trí. Trên 16 ảnh xám 512×512, tốc độ nhúng ròng cực đại trung bình đạt 1,4187 bit trên mỗi điểm ảnh truyền. Tại ER_t=1,0, PSNR trung bình của hai ảnh mang tin đạt khoảng 47,65 dB; toàn bộ 160 phép thử đều cho BER bằng 0 và khôi phục chính xác từng điểm ảnh. Kết quả cho thấy phương pháp đạt dung lượng ổn định, độ méo thấp, bảo vệ tải tin có xác thực và khôi phục chính xác trong khi vẫn sử dụng cơ chế mã hóa ảnh tiêu chuẩn.**

Keywords*— *Reversible data hiding in encrypted images; dual-image RDH; keyed Sudoku coordinate coding; AES-CTR; AES-GCM; adaptive embedding.

Từ khóa*— *Giấu tin thuận nghịch trong ảnh mã hóa; giấu tin thuận nghịch hai ảnh; mã tọa độ Sudoku sinh từ khóa; AES-CTR; AES-GCM; nhúng thích nghi.

I. INTRODUCTION

Digital images are increasingly stored and processed in outsourced environments, including cloud storage, remote medical platforms, digital archives, and collaborative annotation systems. In these settings, embedding additional information into digital content may be required for purposes such as ownership protection, authentication, identification, and metadata management. Digital watermarking provides an established approach to such requirements, with different schemes designed according to the characteristics of the host data and the desired robustness. For example, robust watermarking has been studied for text data using line-spacing

This manuscript was received on July 15, 2026. It was reviewed on August 10, 2026, revised on August 15, 2026 and accepted August 20, 2026.

*Corresponding author

modification [34], while transform-domain techniques based on RDWT, DCT, and SVD have been employed for robust watermarking of color images [35]. However, in applications where the integrity of the host image is as important as the embedded information, robustness alone is not sufficient; the original image may also need to be recovered exactly after the embedded data have been extracted. This requirement motivates reversible data hiding (RDH), which enables the embedded information to be removed while restoring the original content without loss. When the image content must remain confidential during the embedding process, the problem is further extended to reversible data hiding in encrypted images (RDHEI). Research in RDHEI has progressed from early approaches based on modifying encrypted least-significant bits to more advanced strategies involving room reservation, reversible transformations, prediction-based labeling, and high-capacity block coding [6 - 19].

The literature reveals two broad design philosophies. The first creates or preserves redundancy for the data hider. Reserving room before encryption [8], reversible image transformation [9], multi-MSB prediction [10], difference-preserving encryption [11], two-pass PVO [12], adaptive block encoding [13], encrypted adjacency-difference histograms [14], global compression of high bit planes [15], adaptive coding selection [16], asymmetric bit-plane compression [17], and medical-image modulation [18] all exploit a structured carrier. Weng et al. further combined DPE, block-wise difference preservation, and median-preserving PVO [19]. These approaches are technically effective, but they impose an architectural dependency: the encryption stage must preserve or prearrange the exact redundancy required by the embedding stage. The data hider may therefore observe local relations that remain correlated with the plaintext, and changing the cipher generally requires redesigning the hiding rule.

The second philosophy uses two marked images as a joint reversible codeword. Beginning with the two-stegano-image framework [20], researchers have proposed Sudoku-based search [21], center folding [22], complementary coordinate mapping [23], prediction-error shifting [24], and dual-image PVO [25]. The joint state space is larger than that of a single marked image, enabling high payload with exact recovery. Zhang et al. recently proposed Sudoku block mapping (SBM) with adaptive embedding [26]. Their method represents two base-8 digits by two mapped coordinates and controls distortion through a distance threshold. It is elegant and high-capacity, but it is applied to plaintext pixels: the data hider sees the cover image, and the marked outputs remain visually meaningful.

The unresolved issue is therefore architectural. Applying a plaintext RDH rule after conventional encryption is usually infeasible because encryption destroys the correlations on which the inverse rule depends. Preserving those correlations through a specialized transform solves the inversion problem, but ties image protection to the hiding mechanism. A more modular design should use a standard cipher, keep plaintext and the image key hidden from the data hider, reconstruct the ciphertext before decryption, and allow the operating point to follow the requested payload.

The central observation is that Sudoku mapping can be interpreted as a coordinate code rather than as a predictor of natural-image content. Its inverse recovers the input coordinate from the relative displacement of 3×3 blocks. The carrier may therefore be a pair of AES ciphertext bytes rather than a plaintext pixel pair. This interpretation makes it possible to preserve the reversible geometry of Sudoku mapping while separating image encryption from data hiding.

This construction also differs from fixed signed-difference dual-output schemes. Payload symbols are read from Sudoku labels, and the original carrier pair is recovered from block displacement-not from a small difference

alphabet or from max, min, or averaging operations. The encoder evaluates the attainable movement costs and chooses the smallest threshold that can hold the authenticated frame, thereby providing a family of rate-distortion operating points.

The contributions of this work are fourfold:

(1). Keyed Sudoku coordinate coding. A key-derived Sudoku grid, secret pair order, and balanced output orientation encode two base-8 digits in each accepted ciphertext pair without exposing plaintext structure.

(2). Exact ciphertext recovery. The original ciphertext pair is reconstructed uniquely from Sudoku labels and inter-block displacement before AES-CTR decryption; no prediction map or location map is required.

(3). AES-GCM framing. The user payload is framed, encrypted, and authenticated independently of the reversible mapping, so payload confidentiality and integrity do not rely on the Sudoku construction.

(4). Adaptive threshold selection. The smallest movement threshold that accommodates the complete frame is selected, providing explicit rate-distortion control and a measured maximum rate consistent with the analytical ciphertext model.

A same-code plaintext control is used to isolate the effect of moving the coordinate mapping into the encrypted domain. The evaluation reports net payload, PSNR, SSIM, entropy, correlation, NPCR, UACI, BER, and exact recovery.

The remainder of the paper is organized as follows. Section II reviews related RDHEI and dual-image methods and positions the proposed approach. Section III presents the system model and algorithms. Section IV gives a complete embedding and extraction example and establishes reversibility. Section V reports the experiments, and Section VI concludes the paper.

II. RELATED WORK

A. Redundancy-dependent RDHEI

Classical reversible data hiding establishes exact recovery by modifying a compact set of prediction or histogram states. Difference expansion [1] offers a simple reversible rule, but its distortion grows rapidly when large prediction differences are expanded. Histogram shifting [2] limits modification to selected bins, although capacity depends on the availability of dominant peaks and empty bins. Pixel-value ordering (PVO) and its extensions [3 - 5] improve visual quality by embedding only in predictable extrema, but their payload remains tied to local smoothness and ordering. These image-domain assumptions are largely destroyed by conventional encryption, which is why RDHEI methods must either reserve room in advance or deliberately retain a reversible structure after encryption.

Reservation-before-encryption methods [8]–[10] can create substantial room and keep the data hider independent of the plaintext, but they require preprocessing by the content owner and make the embedding capacity depend on prediction or compression efficiency. Difference-preserving and block-based VRAE schemes [11 - 19] move more work to the encrypted domain. Yang et al. [11] retain pixel differences so that encrypted-domain prediction remains possible; Rai et al. [12] and Weng et al. [19] combine preserved ordering with PVO to achieve high fidelity. Gao et al. [13], Anushiadevi and Amirtharajan [14], Yao et al. [15], Xiao et al. [16], Zhang et al. [17], and Chai et al. [18] improve capacity through adaptive block coding, histogram manipulation, bit-plane compression, or medical-image-specific modulation. Their common limitation is architectural coupling: the encryption or preprocessing stage must preserve exactly the redundancy expected by the embedding rule, and auxiliary labels or maps may reduce the effective payload.

Venkatesh et al. [33] represent a recent high-capacity VRAE design based on encrypted pixel differences. In each 2×2 block, three non-reference pixels are rebuilt from an indicator, payload bits, the sign, and the magnitude of their difference from the fourth

pixel. The method is separable and reports 1.1538 bpp on eight images, but the rate varies with the difference class, C5 cases require location information, and rewriting a complete byte can cause considerable carrier-domain distortion. These limitations motivate a reversible mapping whose capacity is less dependent on image content and whose operation does not require a correlation-preserving cipher or a location map.

B. Dual-image reversible coordinate mappings

Dual-image RDH enlarges the reversible state space by representing each payload symbol through two marked outputs [20]. The original two-stegano-image framework demonstrated complementary recovery, but it did not provide explicit rate-distortion control. Sudoku-based search [21] supplies a structured lookup space, although exhaustive or constrained search can increase computational cost. Center folding [22] and multiphase coordinate selection [23] reduce ambiguity, but their capacity and distortion depend on the chosen folding or phase rules. Prediction-error shifting [24] and dual-image PVO [25] improve fidelity by exploiting local prediction, yet they still rely on plaintext redundancy and therefore cannot be transferred directly to a standard ciphertext.

Zhang et al. [26] are the closest methodological precursor. Their Sudoku block mapping places one target in the input pair's center 3×3 block and a second target in a neighboring block selected by the input pair's local position. Two Sudoku labels carry two base-8 digits, while relative block displacement recovers the original pair. The adaptive distance threshold provides an effective rate-distortion mechanism and a theoretical rate of 1.5 bit per transmitted pixel.

The limitation is the operating domain: the original method is applied to plaintext pixels, so the data hider observes the cover and the two marked outputs preserve visible image structure. Applying the same idea after encryption additionally requires authenticated framing, deterministic key synchronization, safe boundary handling, and a proof that the ciphertext pair can still be reconstructed uniquely. These requirements are not addressed by the plaintext formulation.

C. Sudoku in encrypted-image research

Sudoku has also been used in encrypted-image research as a preprocessing or room-reservation aid [31]. Such use demonstrates that Sudoku structures can assist reversible organization, but the Sudoku pattern is not itself the ciphertext-domain carrier code. Qu et al. [32] pursue secure multiparty RDHEI for telemedicine; the protocol strengthens role separation, although its application-specific coordination and lower reported rate make it different from a high-capacity dual-output mapping. Venkatesh et al. [33] operate after encryption and support separable recovery, but their payload depends on encrypted-difference classes and auxiliary C5 handling.

Taken together, the literature leaves a focused gap. Existing high-capacity RDHEI schemes generally depend on preserved redundancy, compressed labels, or image-dependent classes, whereas the most relevant Sudoku dual-image method operates only on plaintext. The proposed method addresses this gap by treating Sudoku as a keyed coordinate code on ordinary AES-CTR ciphertext. It reconstructs the ciphertext before decryption, protects the payload independently with AES-GCM, selects the smallest movement threshold that can carry the authenticated frame, and avoids a location map. The next section develops this construction and proves its reversibility.

D. Positioning of the proposed method and the source of its novelty

The proposed method lies between redundancy-dependent RDHEI and plaintext dual-image mapping. Unlike DPE-, PVO-, and bit-plane-based schemes, it does not require encryption to preserve prediction errors, pixel ordering, or compressible labels. Unlike fixed signed-difference dual-output designs, it does not recover the carrier through arithmetic max, min, or averaging rules. Unlike the original Sudoku block mapping [26], the data hider sees

only AES-CTR ciphertext. Its novelty is the integration of keyed Sudoku coordinate coding, exact ciphertext recovery, AES-GCM framing, and adaptive threshold selection in one map-free architecture.

III. THE PROPOSED METHOD

Figure 1 summarizes the complete data flow. The upper path separates AES-CTR image encryption from payload embedding; the lower path shows that payload extraction and exact ciphertext reconstruction are completed before image decryption. the data hider receives the ciphertext C and the data-hiding key K_h, but never the plaintext O or the image-encryption key K_e. A receiver holding K_h can recover C and the payload; O is available only to a receiver that also possesses K_e .

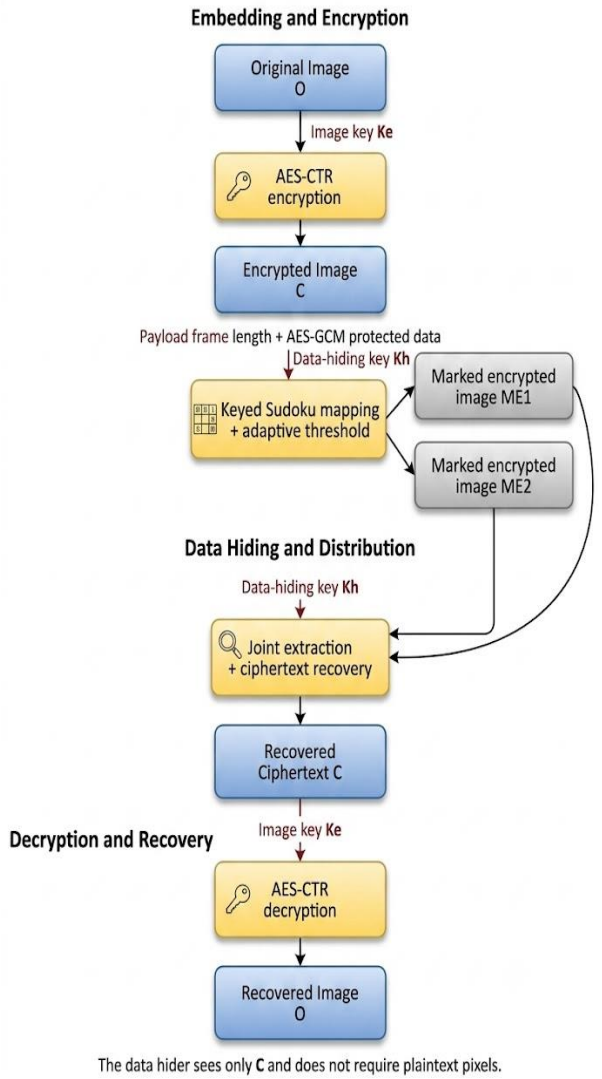

Figure 1. System architecture of the proposed method and separation of image-encryption and data-hiding keys

A. System model and notation

Let $O=\{o_i\}$ be an M×N 8-bit grayscale image, K_e the image-encryption key, K_h the data-hiding key, and S a payload of T bits. The content owner encrypts O with AES-CTR and a unique nonce N_e to obtain C. The data hider receives only C, K_h, and S and produces two marked encrypted images, ME_1 and ME_2. A receiver holding K_h extracts S and reconstructs C; a receiver that additionally holds K_e recovers O. Thus, access to the payload does not imply access to the image.

$$C = AES - CTR(K_e, N_e, O)$$
$$O = AES - CTR(K_e, N_e, C) \quad (1)$$

Net user capacity EC excludes the framing fields. Two rates are reported: ER_c=EC/(MN), normalized by the original image, and ER_t=EC/(2MN), normalized by the two transmitted marked images. (2) By construction, $ER_c=2ER_t$; e.g., the mean maximum operating point ER_t=1.4187 bpp (Section V.C) corresponds to ER_c≈2.8374 bpp with respect to the single original image.

The net user capacity, denoted as EC, excludes the framing fields. Two embedding rates are reported: ER_c=EC/(MN), normalized by the original image size, and ER_t=EC/(MN), normalized by the total size of the two transmitted marked images. Consequently, the relationship between these two rates is expressed as:

$$ER_c = 2ER_t \quad (2)$$

For instance, the mean maximum operating point yields ER_t=1.4187vbpp (corresponding to ER_c = 2.8374 bpp).

B. Authenticated payload framing

The data hider builds a plaintext frame containing the user-bit length and packed payload, then encrypts and authenticates that frame with AES-GCM under K_h. The encrypted-frame length and public GCM nonce are prepended. After padding to a multiple of six bits, the frame is converted into ordered pairs of base-8 digits. This separation is deliberate:

Sudoku mapping provides reversibility, whereas AES-GCM provides payload confidentiality, integrity, and explicit failure detection.

Scope of protection. AES-GCM provides confidentiality and integrity protection for the payload frame only; it does not authenticate the two marked images as transport-level objects. Accordingly, the reported reversibility results assume that ME1 and ME2 are transmitted and received without modification. Any alteration to either marked image may cause ciphertext reconstruction to fail or result in rejection of the extracted payload during GCM authentication. Therefore, AES-GCM should not be regarded as providing end-to-end integrity protection for the pair of marked images.

C. Keyed Sudoku coordinate coding

A valid 9×9 Sudoku grid is derived from K_h by permuting digit labels, bands, rows within bands, stacks, and columns within stacks. Periodic tiling produces a 256×256 reference matrix in which every aligned 3×3 block contains all nine labels exactly once. Raster-adjacent ciphertext bytes form candidate pairs, but K_h determines their processing order. A second key stream supplies one orientation bit per pair and conditionally exchanges the two targets between ME_1 and ME_2, balancing distortion without transmitting side information.

D. Adaptive coordinate embedding

For a ciphertext pair $X_q=(x_q,y_q)$, boundary values in {0,1,2,252,253,254,255} are skipped because a neighboring 3×3 block may extend outside the 8-bit coordinate plane. For an interior pair, the local remainders (x_q mod 3,y_q mod 3) determine one of nine relative blocks. Given the next base-8 digits (d_1,d_2), Y_q is the unique coordinate labeled d_1 in the center block and Z_q is the coordinate labeled d_2 in the selected relative block. A sentinel label resolves the center/equal-digit collision.

$$D_q = \|X_q - Y_q\|_2 + \|X_q - Z_q\|_2 \qquad (3)$$

The encoder examines the finite set of attainable movement costs in ascending order and selects the smallest D_{opt} that can embed the complete authenticated frame. In the final pass, a pair is modified only when $D_q \le D_{opt}$. Boundary, over-threshold, and unused pairs remain equal in ME_1 and ME_2. Neither a location map nor a threshold map is transmitted.

E. Extraction and exact ciphertext recovery

The receiver regenerates the Sudoku matrix, pair order, and orientation bits. Equal marked pairs were not used and are copied directly. For an unequal pair, the receiver reverses the orientation swap and reads d_1 and d_2 from the two Sudoku labels. The block containing the first target identifies the original center block, while the displacement of the second block identifies the original local position of X_q. After the declared frame is collected and verified, the exactly reconstructed ciphertext is decrypted with AES-CTR.

F. Algorithms

Algorithm 1. Adaptive embedding

1) Protect S with AES-GCM; prepend framing fields, pad to six-bit alignment, and convert to base-8 digit pairs.

2) Derive the Sudoku matrix, secret pair order, and output-orientation stream from K_h.

3) For each candidate threshold, scan pairs in key order, skip boundary pairs, compute the two targets and D_q, and count an accepted pair only when D_q is within the threshold.

4) Select the smallest threshold that accommodates the complete framed payload; if no threshold accommodates the complete authenticated frame, report insufficient embedding capacity and abort.

5) Initialize $ME_1=ME_2=C$, rescan accepted pairs, write the two targets, and apply the orientation swap.

6) Output ME_1, ME_2, and the public nonces.

Algorithm 2. Extraction and recovery

1) Regenerate the Sudoku matrix, pair order, and orientation stream.

2) For an equal marked pair, copy the ciphertext pair and append no digits.

3) For an unequal pair, undo orientation, read the two Sudoku labels, derive block displacement, and reconstruct the original ciphertext pair.

4) Parse the encrypted-frame length and stop payload collection after the declared frame while continuing complete ciphertext reconstruction.

5) Authenticate and decrypt the payload with AES-GCM; reject the result if verification fails.

6) Decrypt the exactly reconstructed ciphertext with AES-CTR to recover O.

IV. EMBEDDING AND EXTRACTION EXAMPLE

Consider the ciphertext pair X=(120,121), whose local position in its center 3×3 block is (0,1). Let the payload bits be 001110, corresponding to the base-8 digits $(d_1,d_2)=(1,6)$. In the keyed Sudoku matrix used for this example, label 1 occurs at Y=(120,120) in the center block, and label 6 occurs at Z=(117,121) in the block immediately above. The movement cost is D=1+3=4. With orientation bit zero, ME_1 receives Y and ME_2 receives Z.

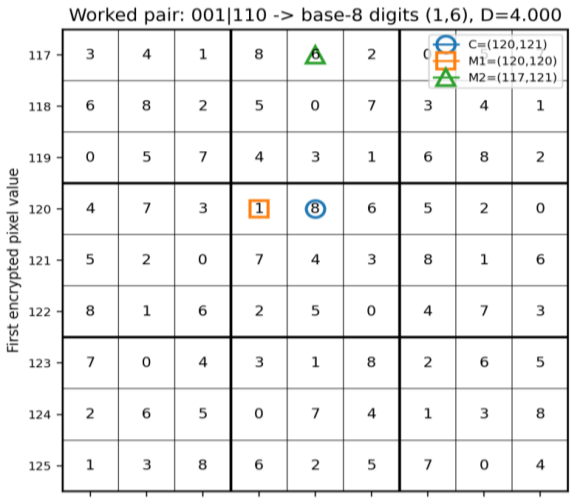

Figure 2. Worked Sudoku-coordinate mapping for ciphertext pair X=(120,121) and payload digits (1,6)

Figure 2 places these coordinates on the tiled Sudoku plane. The first target fixes the original center block; the relative position of the second target's block encodes the local position of X.

The example separates the two functions of the mapping. Labels 1 and 6 carry the six payload bits, whereas the block displacement (-1,0) provides the information needed to restore X=(120,121).

TABLE I. ENCODER AND DECODER STATES FOR THE EXAMPLE IN FIGURE 2

Stage	**Pair / coordinates**	**Decoded information**	**Reason**
Input	X=(120, 121)	Payload 001|110 -> (1,6)	Two base-8 digits per pair
First target	Y=(120, 120)	Sudoku label 1	Y remains in original center block
Second target	Z=(117, 121)	Sudoku label 6	Z lies in the block selected by local position (0,1)
Marked outputs	ME_1=Y, ME_2=Z	Orientation bit 0	Orientation bit 1 would swap outputs
Extraction	Labels 1 and 6	Bits 001110	Read labels after undoing orientation
Recovery	Block displacement (-1,0)	Local position (0,1)	Recover center-block coordinate X=(120,121)
Final	Recovered C byte pair	AES-GCM verifies payload; AES-CTR restores image	Exact ciphertext recovery precedes decryption

Table I presents the same example as a sequence of encoder and decoder states, distinguishing the quantities that carry payload from those that guarantee reversibility.

The final row makes the processing order explicit: the ciphertext is reconstructed first; payload authentication and image decryption follow.

Reversibility argument

Proposition 1. Under the correct keys and intact marked outputs, every embedded ciphertext pair is recovered uniquely. The first mapped coordinate lies in the input pair's center block, and the second lies in the relative block selected by the input pair's local position. Their block displacement therefore determines that local position. Because each 3×3 block contains every Sudoku label exactly once, the two payload digits are also unique. The sentinel resolves the center/equal-digit ambiguity, while equal outputs identify unused pairs. Consequently, the ciphertext is reconstructed

byte for byte and AES-CTR decryption recovers O exactly.

V. EXPERIMENTAL EVALUATION

A. Setup and test images

The implementation uses Python, NumPy, Pillow, scikit-image, and the cryptography package. Sixteen 8-bit grayscale BMP images of size 512×512 were evaluated: Airplane, Barbara, Blob, Boat, Camera, Cameraman, Car, Coins, Goldhill, Moon, Page, Sailboat, cabeza, chest, forest, lena, manoeuvre, pepper, phone, and shirt. Deterministic experimental keys and unique per-image nonces were derived from a master seed. Payloads were tested at ER_t=0.2, 0.5, 1.0, and at the image-dependent maximum.

Figure 3 shows sixteen representative images from the sixteen-image collection. The subset includes smooth, textured, natural, medical, and technical content and is used only to illustrate the diversity of the plaintext inputs. AES-CTR removes the visible texture of these inputs before embedding. The maximum ciphertext-domain rate should therefore vary less with image content than the rate of a prediction-based plaintext method; Section V-C tests this expectation. Two controlled pipelines were evaluated under the same implementation. The proposed method applies the keyed Sudoku coordinate code to AES-CTR ciphertext and protects the payload with AES-GCM. The Sudoku control applies the same pair order, framing, orientation balancing, and threshold search directly to plaintext. It is an ablation for isolating the effect of the carrier domain, not a byte-exact reproduction of Zhang et al. [26].

Figure 3. The test images

A visual check on Airplane confirmed that the AES-CTR carrier and both marked outputs were noise-like and that the recovered image was pixel-identical to the original. Because exact recovery is already verified quantitatively, a redundant image montage is not included.

B. Capacity and image quality

Table II reports the sixteen-image averages at four operating points. Capacity is normalized by the two transmitted outputs through ER_t, while PSNR and SSIM are measured between each carrier and its corresponding marked output.

TABLE II. MEAN NET EMBEDDING RATE AND CARRIER-TO-MARKED QUALITY OVER SIXTEEN IMAGES

Method	Target ER_t	EC (bits)	Net ER_t	PSNR1	PSNR2	SSIM1	SSIM2
Sudoku control	0.2	104,856.0	0.199997	58.533	58.529	0.999309	0.999309
	0.5	262,144.0	0.500000	52.815	52.820	0.997476	0.997474
	1.0	524,288.0	1.000000	47.837	47.829	0.992119	0.992105
	max	782,224.8	1.491975	44.465	44.463	0.982865	0.982839
The proposed method	0.2	104,856.0	0.199997	58.503	58.490	0.999991	0.999991
	0.5	262,144.0	0.500000	52.711	52.718	0.999967	0.999967
	1.0	524,288.0	1.000000	47.651	47.657	0.999895	0.999896
	max	743,794.8	1.418676	44.697	44.691	0.999794	0.999793

The maximum-rate rows show the principal trade-off. The plaintext control can use almost every coordinate pair, whereas the proposed method loses about 5.4% of pairs at ciphertext boundaries. In exchange, the data hider does not see the plaintext, and the marked outputs retain ciphertext-like statistics.

At matched rates, transferring the same coordinate code from plaintext to AES ciphertext causes only a small PSNR change. At ER_t=1.0, the proposed outputs reach 47.651 and 47.657 dB, compared with 47.837 and 47.829 dB for the Sudoku control. The orientation mechanism keeps the average PSNR difference between ME_1 and ME_2 below 0.02 dB. The high ciphertext-domain SSIM values describe numerical fidelity to C and should not be read as perceptual similarity to O.

Figure 4 shows how PSNR and SSIM vary with the transmission-normalized payload. Both curves decline smoothly as the adaptive threshold admits more pairs. The absence of an abrupt quality break supports D_{opt} as a practical rate-distortion control rather than a fixed embedding switch

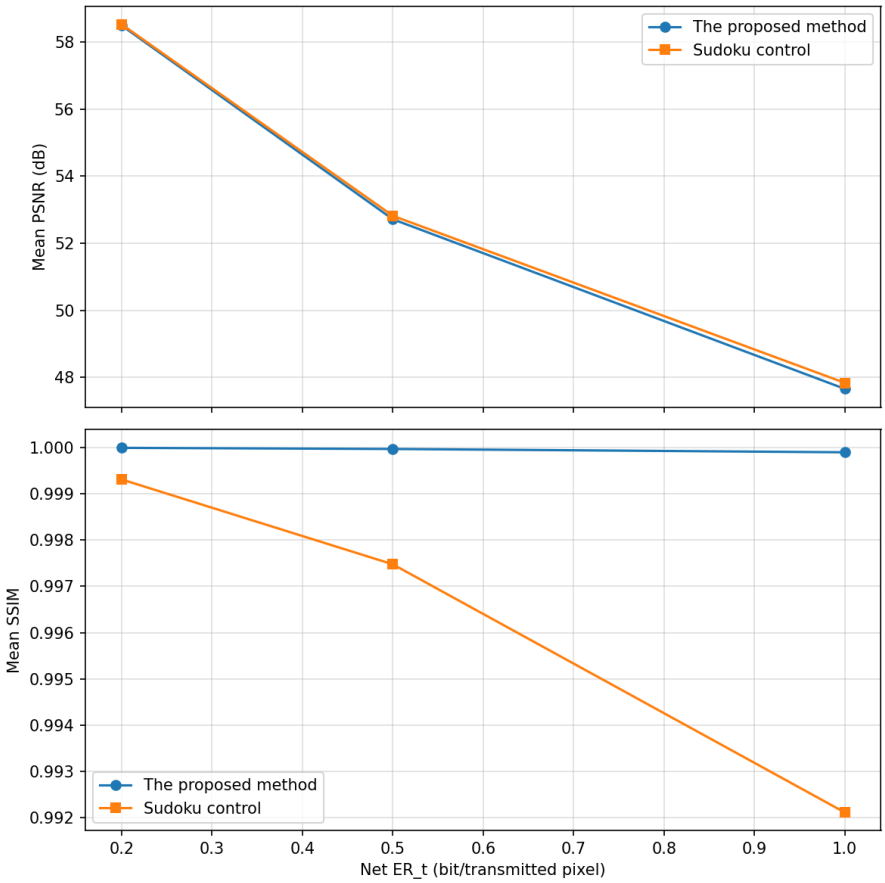

Figure 4. Mean carrier-to-marked PSNR and SSIM versus transmission-normalized payload

C. Analytical and measured maximum rate

Under AES-CTR, ciphertext bytes are approximately uniform. A byte is admissible with probability 249/256 and a pair is admissible with probability $(249/256)^2 \approx 0.9461$. Since each accepted pair carries six bits, the predicted maximum rate is:

$$E[ER_t,max] = 1.5(249/256)^2 \approx 1.4192 \text{ bpp}. \quad (4).$$

The measured maximum ER_t ranges from 1.4169 to 1.4201 bpp, with a mean of 1.4188 bpp and a 95% confidence interval of ±0.00038. Its difference from the analytical prediction is only 0.0004 bpp, indicating that the limiting factors are the ciphertext boundary rule and framing overhead rather than plaintext texture.

Table III compares the closed-form prediction with the measurements. The model

depends only on the admissible-byte probability and the six payload bits carried by each accepted pair. The close agreement explains the narrow rate variation across the sixteen images: after AES-CTR encryption, capacity is controlled primarily by the byte boundary rule rather than by the smoothness or texture of the original image.

TABLE III. ANALYTICAL PREDICTION AND MEASURED MAXIMUM EMBEDDING RATE

Quantity	Analytical value	Measured result	Interpretation
Valid-pair fraction	0.9461	≈0.9459	Boundary probability of uniform ciphertext
Maximum ER_t	1.4192 bpp	1.4187 bpp	Predictable, texture-independent capacity
Mean maximum payload	≈744 kbit	743,794.8 bit	Net authenticated payload
BER / recovery error	0	0	Exact payload and image recovery

D. Comparative protocol and reproducibility boundaries

Internally measured results are kept separate from values reported in earlier work. The proposed method and the Sudoku control were run on all sixteen images and four payload levels. The published rate-PSNR curve of Zhang et al. is used only as a plaintext-Sudoku reference. For Venkatesh et al. [33], the disclosed embedding and recovery rules were implemented and validated against their C1 example. Because the paper does not fully specify the encryption recurrence or the compression cost of the location map, reproduced and published values are reported separately. This distinction prevents implementation assumptions from being presented as results of the original authors.

For clarity, three quantities should be distinguished when interpreting the results of Venkatesh et al. [33]: (i) the embedding rates reported by the original authors; (ii) the gross payload obtained from our independent implementation of the disclosed embedding and recovery procedures; and (iii) the conservative reproduced net payload after accounting for location information. For the conservative case reported in Table IV, we deduct one uncompressed location bit for each non-reference pixel, corresponding to an overhead of 3MN/4 bits for an M×N image. Since [33] does not provide sufficient details on the compression of the location map to enable byte-exact reproduction, this deduction is adopted as an explicit lower-bound assumption in our reproduction and should not be attributed to the original authors.

E. Embedding-capacity comparison

Table IV compares the two implementations on sixteen images using bits per transmitted pixel (bptp). For the proposed two-output method, bptp=EC/(2MN)=ER_t. Venkatesh et al. produce one stego image, so their bpp is numerically equal to bptp. The reproduced Venkatesh values are conservative net rates obtained after deducting one uncompressed location bit per non-reference pixel. Under this accounting, the proposed method averages 1.4187 bptp, compared with 0.8053 bptp for the reproduced Venkatesh layer.

TABLE IV. ACTUAL MAXIMUM NET EMBEDDING RATE ON SIXTEEN IMAGES IN BITS PER TRANSMITTED PIXEL (BPTP)

Image	Venkatesh reproduction Net bptp	The proposed method Net bptp
Airplane	1.0029	1.4176
Barbara	0.6504	1.4195
Blob	1.1395	1.4185
Boat	0.6442	1.4190
Cameraman	1.1048	1.4183
Car	0.7628	1.4188
Goldhill	0.6700	1.4206
Sailboat	0.5896	1.4190
Cabeza	1.3974	1.4183
Chest	1.0957	1.4181
Forest	0.4554	1.4193
Lena	0.8935	1.4173
Manoeuvre	0.4231	1.4179
Pepper	0.8255	1.4189
Phone	0.9242	1.4209
Shirt	0.3052	1.4173
Mean	**0.8053**	**1.4187**

The proposed rate remains close to 1.4187 bptp for every image because admissibility is governed mainly by the AES-ciphertext boundary probability. The reproduced Venkatesh rate varies more widely because its payload depends on image-specific difference classes and C5 location overhead.

F. Marked-image quality and exact recovery

At ER_t=0.2, 0.5, and 1.0, the proposed method and the Sudoku control exhibit similar PSNR trends. At maximum payload, the proposed method obtains about 44.69 dB between the AES carrier and each marked output. The reproduced Venkatesh layer gives 8.81 dB between its carrier and stego image because each non-reference pixel is rebuilt as an indicator-payload-sign-magnitude codeword. The infinite PSNR reported by Venkatesh et al. refers to the exactly recovered original image, not to the fidelity of the marked encrypted image.

All reproduced Venkatesh runs recover the carrier and the extracted payload prefix without error under the documented per-pixel C5 flags. The proposed method likewise yields zero BER and exact ciphertext and plaintext recovery in every trial, including the S=0 case. Both constructions are therefore reversible in the tested noise-free setting, but their carrier-domain distortion differs substantially.

TABLE V. REFERENCE-ONLY, RATE-MATCHED PSNR COMPARISON WITH THE PLAINTEXT SUDOKU METHOD IN [26] UNDER DIFFERENT CARRIER DOMAINS

ER_t	The proposed method PSNR1	The proposed method PSNR2	Published Sudoku PSNR	Interpretation
0.2	58.503	58.490	58.80	Same rate; different carrier domains
0.5	52.711	52.718	53.05	Same rate; different carrier domains
1.0	47.651	47.657	48.20	Same rate; different carrier domains

Table V compares only payload rates reported by both studies. The published Sudoku values measure plaintext cover-to-marked quality, whereas the proposed values measure AES-ciphertext carrier-to-marked quality. The comparison therefore concerns the rate-distortion trend, not perceptual equivalence. At each rate, the plaintext predecessor has slightly higher PSNR, as expected from its content-adaptive carrier. The proposed method accepts this modest penalty in exchange for hiding the plaintext from the data hider and retaining standard encryption.

G. Security-related statistical comparison

At maximum payload, the marked outputs of the proposed method retain ciphertext-like statistics, with mean entropy of 7.9984 bits and horizontal correlation close to zero. The reproduced Venkatesh layer shows lower post-embedding entropy because its indicator codewords occupy a more restricted value distribution. The original paper reports entropy near 7.9993 bits, but that value cannot be reproduced without the complete encryption specification. Entropy and correlation are statistical diagnostics, not proofs of cryptographic security.

NPCR and UACI are likewise descriptive and depend on the image pairs used in the calculation. Security in the proposed method rests on standard AES assumptions, key separation, and nonce discipline—not on the Sudoku mapping. In deployment, fresh AES-CTR and AES-GCM nonces must be used for every key session, and the same key-nonce pair must never be reused.

Paired-output security limitation. An adversary may jointly observe ME1 and ME2 and exploit statistical dependencies between the two outputs introduced by the reversible mapping. The entropy, adjacent-pixel correlation, NPCR, and UACI results reported in this work characterize their marginal statistical properties only; they do not constitute a comprehensive steganalysis evaluation and should not be interpreted as evidence of resistance to detection based on the joint observation of the paired outputs. Accordingly, dedicated paired-output steganalysis remains an important direction for future evaluation.

TABLE VI. CAPACITY AND SECURITY-RELATED STATISTICAL DIAGNOSTICS UNDER THE REPORTED PROTOCOLS

Method	Capacity (bpp)	Entropy after hiding	Horizontal correlation	NPCR	UACI
The proposed method	1.4187	7.998444	0.000324	99.6077	30.6410
Venkatesh et al.	1.3845	7.999275	-0.000275	99.6414	33.5669

Table VI gathers the diagnostics available under sufficiently similar reporting conventions. Entropy and adjacent-pixel correlation are the most directly comparable. NPCR and UACI are retained for completeness, but differences in the underlying protocols preclude a direct ranking.

Both marked-image sets have entropy close to eight bits and near-zero correlation. The difference between 7.998444 and 7.999275 is too small to support a meaningful security ranking; it merely shows that both distributions are close to uniform.

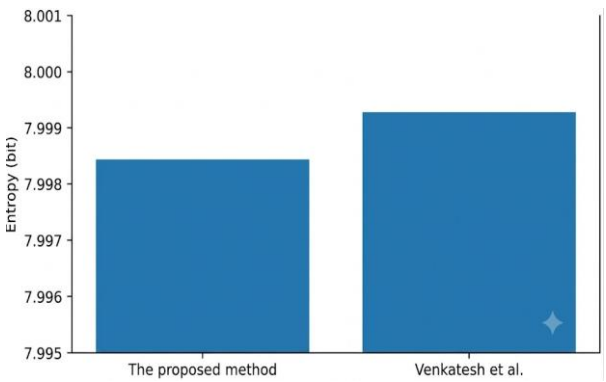

Figure 5. Post-embedding entropy on the four image names common to both studies

Figure 5 magnifies the entropy values on the common subset; the truncated vertical axis is used only to make the small numerical difference visible.

The slightly higher Venkatesh bar should not be interpreted as stronger encryption. The proposed method's security argument rests on AES-CTR, AES-GCM, nonce discipline, and key separation rather than on a marginal entropy difference.

VI. CONCLUSION

This paper introduced a secure dual-image RDHEI architecture built around four complementary ideas: keyed Sudoku coordinate coding, exact ciphertext recovery, AES-GCM-protected framing, and adaptive threshold selection. AES-CTR protects the image independently of the hiding layer; Sudoku labels carry the payload digits; inter-block displacement reconstructs each ciphertext pair before decryption; and D_{opt} selects the lowest-cost mappings that can hold the complete authenticated frame. On sixteen 512×512 images, the method achieved a mean maximum net rate of 1.4187 bptp. At ER_t=1.0, the marked outputs averaged about 47.65 dB, and every tested payload and image was recovered exactly. The comparison with plaintext Sudoku mapping and the independently reproduced Venkatesh layer clarifies the method's operating advantage without conflating incompatible protocols: predictable net capacity, low carrier-domain distortion, standard cryptographic protection, and map-free recovery. Future work should address paired-output steganalysis and reduce the transmission cost of the two-image channel.

FUNDING STATEMENT

This research is funded by University of Transport and Communications (UTC) under grant number T2026-CN-006.

REFERENCES

[1] J. Tian, "Reversible data embedding using a difference expansion", *IEEE Trans. Circuits Syst. Video Technology*, vol. 13, no. 8, pp. 890–896, 2003, DOI: 10.1109/TCSVT.2003.815962.

[2] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding", *IEEE Trans. Circuits Syst. Video Technology*, vol. 16, no. 3, pp. 354–362, 2006, DOI: 10.1109/TCSVT.2006.869964.

[3] X. Li, J. Li, B. Li, and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion", *Signal Process*, vol. 93, no. 1, pp. 198–205, 2013, DOI: 10.1016/j.sigpro.2012.07.025.

[4] B. Ou, X. Li, Y. Zhao, R. Ni, and Y. Q. Shi, "Pairwise prediction-error expansion for efficient reversible data hiding", *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5010–5021, 2013, DOI: 10.1109/TIP.2013.2281422.

[5] C. Y. Weng, H. Y. Weng, and C. T. Huang, "High-fidelity reversible data hiding based on PVO and median preserving", *J. Supercomput.*, vol. 78, pp. 8367–8388, 2022, DOI: 10.1007/s11227-021-04226-0.

[6] X. Zhang, "Reversible data hiding in encrypted image", *IEEE Signal Process. Lett.,* vol. 18, no. 4, pp. 255–258, 2011, DOI: 10.1109/LSP.2011.2114651.

[7] X. Zhang, "Separable reversible data hiding in encrypted image", *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 826–832, 2012.

[8] K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, "Reversible data hiding in encrypted images by reserving room before encryption", *IEEE Trans. Inf. Forensics Security,* vol. 8, no. 3, pp. 553–562, 2013, DOI: 10.1109/TIFS.2013.2248725.

[9] W. Zhang, H. Wang, D. Hou, and N. Yu, "Reversible data hiding in encrypted images by reversible image transformation", *IEEE Trans. Multimedia*, vol. 18, no. 8, pp. 1469–1479, 2016.

[10] P. Puteaux and W. Puech, "An efficient MSB prediction-based method for high-capacity reversible data hiding in encrypted images", *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 7, pp. 1670–1681, 2018, DOI: 10.1109/TIFS.2018.2799381.

[11] C. H. Yang, C. Y. Weng, and J. Y. Chen, "High-fidelity reversible data hiding in encrypted image based on difference-preserving encryption", *Soft Comput.*, vol. 26, pp. 1727–1742, 2022, DOI: 10.1007/s00500-022-06745-1.

[12] A. K. Rai, H. Om, S. Chand, and S. Agarwal, "Reversible data hiding in encrypted image using two-pass pixel value ordering", *J. Inf. Secur. Appl*., vol. 76, art. 103545, 2023, DOI: 10.1016/j.jisa.2023.103545.

[13] K. Gao, J. H. Horng, and C. C. Chang, "High-capacity reversible data hiding in encrypted images based on adaptive block encoding", *J. Vis. Commun. Image Represent*., vol. 84, art. 103481, 2022, DOI: 10.1016/j.jvcir.2022.103481.

[14] R. Anushiadevi and R. Amirtharajan, "Separable reversible data hiding in an encrypted image using the adjacency pixel difference histogram", *J. Inf. Secur. Appl*., vol. 72, art. 103407, 2023, DOI: 10.1016/j.jisa.2022.103407.

[15] Y. Yao, K. Wang, Q. Chang, and S. Weng, "Reversible data hiding in encrypted images using global compression of zero-valued high bitplanes and block rearrangement", *IEEE Trans. Multimedia*, vol. 26, pp. 3701–3714, 2024, DOI: 10.1109/TMM.2023.3314975.

[16] F. Xiao, K. Wang, Y. Shen, and Y. Yao, "High-capacity reversible data hiding in encrypted images based on adaptive block coding selection", *J. Vis. Commun. Image Represent.,* vol. 104, art. 104291, 2024, DOI: 10.1016/j.jvcir.2024.104291.

[17] X. Zhang, F. He, C. Yu, X. Zhang, C. N. Yang, and Z. Tang, "Reversible data hiding in encrypted images with asymmetric coding and bitplane block compression", *IEEE Trans. Multimedia*, vol. 26, pp. 10174–10188, 2024, DOI: 10.1109/TMM.2024.3405717.

[18] X. Chai, G. Cao, Z. Fu, Z. Gan, B. Wang, and Y. Zhang, "High-capacity reversible data hiding in encrypted medical images using adaptive pixel-modulation and HBP-RMC", *Biomed. Signal Process. Control,* vol. 95, art. 106424, 2024, DOI: 10.1016/j.bspc.2024.106424.

[19] C. Y. Weng, T. W. Lin, H. Y. Weng, and C. T. Huang, "Enhanced reversible data hiding in encrypted images via median preserving pixel value ordering and block-wise difference preservation", *Signal Image Video Process*., vol. 19, art. 424, 2025, DOI: 10.1007/s11760-025-04054-2.

[20] C. C. Chang, T. D. Kieu, and Y. C. Chou, "Reversible data hiding scheme using two steganographic images," *Proc. IEEE TENCON*, Taipei, Taiwan, 2007, pp. 1–4, DOI: 10.1109/TENCON.2007.4483783.

[21] T. S. Nguyen, C. C. Chang, and T. F. Chung, "A reversible data hiding scheme based on the Sudoku technique", *J. Syst. Softw*., vol. 102, pp. 64–74, 2015, DOI: 10.1016/j.jss.2014.12.029.

[22] T. C. Lu, J. H. Wu, and C. C. Huang, "Dual-image-based reversible data hiding method using center folding strategy", *Signal Process.,* vol. 115, pp. 195–213, 2015.

[23] I. F. Jafar, K. A. Darabkh, R. T. Al-Zubi, and R. R. Saifan, "An efficient reversible data hiding algorithm using two steganographic images", *Signal Process.,* vol. 128, pp. 98–109, 2016.

[24] H. Yao, F. Mao, Z. Tang, and C. Qin, "High-fidelity dual-image reversible data hiding via prediction-error shift", *Signal Process.,* vol. 170, art. 107447, 2020.

[25] Y. Niu and S. Shen, "A novel pixel value ordering reversible data hiding based on dual-image", *Multimedia Tools Appl*., vol. 81, pp. 13751–13771, 2022, DOI: 10.1007/s11042-022-12149-y.

[26] Y. Zhang, Y. Yao, C. C. Lin, and C. C. Chang, "Dual-image reversible data hiding based on Sudoku block mapping and adaptive embedding", *J. Internet Technol.,* vol. 26, no. 2,

pp. 147–163, 2025, DOI: 10.70003/160792642025032602001.

[27] A. Abba, J. S. Teh, and M. Alawida, "Towards accurate keyspace analysis of chaos-based image ciphers", *Multimedia Tools Appl.,* vol. 83, pp. 79047–79066, 2024, DOI: 10.1007/s11042-024-18628-8.

[28] National Institute of Standards and Technology, "Advanced Encryption Standard (AES)", *FIPS PUB 197-upd1*, 2023, DOI: 10.6028/NIST.FIPS.197-upd1.

[29] M. Dworkin, "Recommendation for block cipher modes of operation: Methods and techniques", NIST SP 800-38A, 2001, DOI: 10.6028/NIST.SP.800-38A.

[30] M. Dworkin, "Recommendation for block cipher modes of operation: Galois/Counter Mode (GCM) and GMAC," NIST SP 800-38D, 2007, DOI: 10.6028/NIST.SP.800-38D.

[31] X.-Y. Li, X.-B. Zhou, Q.-L. Zhou, S.-J. Han, and Z. Liu, "High-capacity reversible data hiding in encrypted images by information preprocessing", *Complexity*, vol. 2020, Art. no. 6989452, 2020, DOI: 10.1155/2020/6989452.

[32] L. Qu, M. Li, and P. Chen, "Reversible data hiding in encrypted image with secure multi-party for telemedicine applications", *Biomed. Signal Process. Control*, vol. 93, Art. no. 106209, 2024, DOI: 10.1016/j.bspc.2024.106209.

[33] V. Venkatesh, R. Anushiadevi, P. V. Meikandan, H. Mahalingam, and R. Amirtharajan, "Separable reversible data hiding by vacating room after encryption using encrypted pixel difference", *Sci. Rep.,* vol. 15, Art. no. 11916, 2025, DOI: 10.1038/s41598-025-96152-x.

[34] A. Kozachok and S. Kopylov, "Robust text watermarking based on line shifting," Journal of Science and Technology on Information Security,vol. 1, no. 7, pp. 3–13, 2020, doi: 10.54654/isj.v7i1.50.

[35] V. T. Son and T. M. Duc, "A robust watermarking method based on RDWT-DCT-SVD in the YCbCr color space," Journal of Science and Technology on Information Security, vol. 3, no. 23, pp. 5–14, 2024,doi: 10.54654/isj.v3i23.1054..

ABOUT THE AUTHOR

Cao Thi Luyen

Workplace: Faculty of Information Technology, University of Transport and Communications, Hanoi, Viet Nam.

Email: luyenct@utc.edu.vn

Education: B.S. in Mathematics and Computer Science, VNU University of Science, 2001; M.S. in Mathematical Assurance for Computing Systems, VNU University of Science, 2005; Ph.D. in Mathematical Foundations for Information Technology, Academy of Military Science and Technology, 2018.

Research interests: information security, digital watermarking, reversible data hiding, and encrypted-image processing.

Tên tác giả: **Cao Thị Luyên**

Cơ quan công tác: Khoa Công nghệ Thông tin, Trường Đại học Giao thông Vận tải, Hà Nội.

Email: luyenct@utc.edu.vn

Quá trình đào tạo: Nhận bằng cử nhân Toán cơ–Tin học và thạc sỹ đảm bảo toán học do Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội cấp lần lượt ở các năm 2001, 2005; Tiến sĩ Cơ sở toán học cho tin học, Viện Khoa học và Công nghệ quân sự, năm 2018.

Hướng nghiên cứu: an toàn thông tin, thủy vân số, giấu tin thuận nghịch và xử lý ảnh mã hóa.