scientific reports

Check for updates

OPEN

A multi-branch feature enhancement-based detection and hierarchical chaotic encryption fusion method for sensitive targets in remote sensing images

Qingqi Zhang[1], Hao Wang[1]✉, Xiantao Li[2], Shitao Zhang[3] & Jinzhou Liu[4]

Remote sensing images used in military reconnaissance contain a large amount of sensitive information, and any leakage may pose a serious threat to national security. To address the need for high-precision detection of sensitive targets and to mitigate information leakage risks, this study proposes a remote sensing image processing framework that integrates multi-object detection with hierarchical chaotic encryption. Based on the YOLOv7-tiny architecture, a Multi-branch Enhanced Feature Aggregation Block (MEFABlock) is designed, which incorporates multi-scale convolutions and attention mechanisms to effectively enhance feature extraction in complex remote sensing scenes. In addition, a coordinate convolution module is introduced before the detection head to strengthen spatial position modeling, thereby achieving higher detection accuracy while maintaining a lightweight network structure.In the encryption stage, a novel two-dimensional chaotic system is constructed, and an improved hash-based method is proposed to generate the initial parameters of the chaotic system. The Josephus-ring permutation process is further enhanced, and a three-stage diffusion–permutation–diffusion encryption structure is employed. The improved detection network is first used to precisely locate sensitive regions in the image; the regions inside the detected bounding boxes are then locally encrypted and embedded back into the original image to achieve information concealment. Finally, global encryption is performed using the chaotic system to ensure end-to-end data security.Experiments conducted on the publicly available MAR20 military aircraft remote sensing dataset demonstrate that the proposed method improves Precision (P) by 2.4%, Recall (R) by 2.7%, mAP@0.5 by 2.7%, and mAP@0.5:0.95 by 2.2% compared with the baseline model. The encrypted remote sensing images achieve an information entropy of 7.9997, and meet high security standards in key metrics such as NPCR and UACI. Overall, the proposed encryption framework achieves high target detection accuracy and strong information protection performance, exhibiting robust potential for practical engineering applications.

Keywords Multi-object detection, YOLO, Feature enhancement, Attention mechanism, Chaotic encryption, Sensitive information protection

In modern warfare and intelligent combat systems, accurately detecting battlefield-sensitive targets plays a crucial role in enhancing situational awareness, supporting reconnaissance operations, and assisting in firepower allocation[1]. With the widespread application of high-resolution satellite remote sensing imagery, achieving efficient detection of sensitive targets while ensuring information security has become a critical challenge in the field of military image processing.

In recent years, with the rapid development of deep learning technologies, convolutional neural network (CNN)-based object detection methods have been widely applied in image analysis and processing tasks. Currently, CNN-based object detection algorithms can be categorized into two-stage and one-stage methods[2].

[1]Electronic Information Engineering College, Changchun University, Changchun 130022, China. [2]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China. [3]Wuhan Second Ship Design and Research Institute, Wuhan 430000, China. [4]School of Electrical and Computer Engineering, Jilin Jianzhu University, Changchun 130119, China. ✉email: blueswang@ccu.edu.cn

Common two-stage algorithms include Faster R-CNN[3] and R-FCN[4]. Although two-stage algorithms offer high detection accuracy, their complex structures and slower inference speeds make them difficult to deploy in real-time processing and edge devices, which is particularly limiting in scenarios such as remote sensing monitoring and aerial reconnaissance where detection speed and efficiency are critical.

Compared with two-stage algorithms, one-stage methods, represented by the YOLO (You Only Look Once) series from YOLOv5 onwards[5–13], have gradually become mainstream detection solutions in scenarios such as remote sensing, traffic monitoring, and security surveillance due to their good detection performance and real-time inference capability. For remote sensing object detection, Wang et al.[14] proposed the AG-YOLO model guided by attention mechanisms, achieving efficient object detection in remote sensing scenarios by introducing attention mechanisms and rotation box parameters in the detection head. Wu et al.[15], based on YOLOv8, improved the recognition accuracy of targets in remote sensing images by integrating the EMA attention mechanism and the WIoU loss function. Although these studies contributed to detection accuracy under remote sensing backgrounds, their selected targets were not military objects, limiting their applicability to military reconnaissance.

For the detection of military vehicles in different scenarios, some scholars have conducted research on YOLO algorithms. Hong et al.[16] modified the loss function of YOLOv3-tiny and used the K-means clustering algorithm to optimize anchor boxes, improving detection accuracy for tanks, infantry fighting vehicles, and other units. Liu et al.[17] proposed the improved MOKP-YOLO algorithm based on YOLOv8n, enhancing detection performance for military targets from drone perspectives through the design of a unified detector and a key component feature integration module.

Military aircraft are key aerial assets in modern warfare and national security defense, characterized by complex maneuverability and diverse tactical applications. Identifying and classifying military aircraft in remote sensing imagery has significant practical value and strategic importance. Xi et al.[18] proposed MPS-YOLO, a multi-scale information fusion network based on PixelShuffle and YOLO, which improved the recognition accuracy of critical targets such as military aircraft in remote sensing scenarios; however, their algorithm required substantial computational resources. Liu et al.[19] introduced the TripletAttention module and an efficient dynamic upsampler into the YOLOv8s baseline model, enabling better detail capture in images, but there remains room for further improvement in detection accuracy.

In addition to YOLO-based methods, researchers have also explored alternative algorithms. Xi et al.[20] proposed a novel single-order structure-adaptive object detection (SOOD) network, introducing a new rotation angle encoder (RAE) along with structure-adaptive label assignment (SALA) and structure-adaptive confidence estimation (SACE) to more accurately localize object positions. Shi et al.[21] proposed a novel anchor-free detection network that leverages point set representation, integrating a progressive class-aware dual-branch module (PCA-DB) and an instance-guided enhancement module (IGEM). These models have achieved favorable detection results in complex backgrounds. However, dedicated optimizations for military aircraft categories remain relatively limited, and there is still a performance gap compared to YOLO-based approaches in balancing real-time performance and accuracy.

While high-precision object detection methods enable the acquisition of sensitive information in images, they also raise concerns about the security of this information. In particular, in military remote sensing imagery, unauthorized access or malicious tampering of sensitive targets such as military aircraft can result in severe information leakage and security threats. Therefore, it is imperative to integrate efficient object detection with information protection technologies, establishing an image processing framework that balances detection performance and data security.

Chaotic encryption has become an important research direction in the field of information security. Owing to its desirable properties—such as high sensitivity to initial conditions, controllable parameters, and intrinsically complex dynamic behavior—it has been widely applied to image encryption and information hiding tasks. With the continuous development of computing technologies, emerging paradigms such as neural networks and bio-inspired computation have also been introduced into chaos-based cryptographic systems. As chaotic systems exhibit strong unpredictability and dynamical complexity, they are well suited for secure image encryption applications[22]. Li et al.[23] proposed a structurally simple two-dimensional enhanced logistic modular map and further designed a vector-level chaotic image encryption algorithm based on this map, achieving promising encryption performance. However, exponential-type chaotic maps are prone to numerical degradation due to floating-point errors during digital implementation. To address this, Yu et al.[24] constructed a six-dimensional hyperchaotic system by incorporating a memristor into a five-dimensional chaotic framework, demonstrating superior hyperchaotic characteristics.In the domain of chaotic image encryption, the integration of memristors with neural networks has also attracted increasing attention. Yu et al.[25]– [26] introduced memristive devices into Hopfield neural networks to enhance their dynamical properties, and further implemented an image encryption circuit using FPGA technology. Ye et al.[27] employed discrete wavelet transform for image preprocessing and utilized cellular automata and other computational mechanisms to achieve secure image encryption.

In the field of chaotic encryption, researchers have proposed various novel encryption schemes. Wang et al.[28] designed a block-wise Arnold scrambling and bit-level permutation scrambling method, combined with a dual diffusion method involving XOR operations between a chaotic sequence and a secondary hash index chain, thereby enhancing the anti-attack capability of remote sensing images. Teng et al.[29] encrypted facial information using a combination of DNA diffusion and Fisher-Yates shuffling, effectively protecting facial privacy. Kumar et al.[30] proposed a new hybrid image encryption method combining chaotic maps and genetic algorithms, which improved encryption security, complexity, and robustness against various attacks. Ammar Odeh et al.[31] presented a lightweight secure image encryption algorithm based on the Tent map chaotic system, which achieved high encryption strength while maintaining computational efficiency. This lightweight method effectively obscures image content and offers good robustness. Alenrex Maity et al.[32] proposed an encryption framework integrating

a 5D hyperchaotic system, wavelet lifting transform, and Burrows-Wheeler transform, providing enhanced protection for images containing sensitive data. Zhang et al.[33] proposed a novel image encryption approach based on a self-developed chaotic system and DNA coding, introducing a new decimal-to-binary conversion method; however, the chaotic system was relatively complex.

Recent studies have demonstrated that many chaotic image encryption schemes still exhibit significant security vulnerabilities when evaluated under practical attack models. In particular, several recent advances in cryptanalysis have revealed that numerous chaotic encryption algorithms—despite their sophisticated designs and seemingly strong statistical performance—can still be successfully compromised through known-plaintext attacks, chosen-plaintext attacks, or structural analysis.

For example, Feng et al.[34] conducted a comprehensive cryptanalysis of the image encryption scheme based on a Feistel network and dynamic DNA encoding (IES-FD). Their analysis revealed severe weaknesses in key scheduling, Hill encryption, DNA operations, and the diffusion mechanism, and proposed a chosen-plaintext attack capable of fully recovering the plaintext, indicating that the scheme fails to meet practical security requirements. Wen et al.[35] examined the IEC-BPMC algorithm, which is based on binary bit-plane extraction and multiple chaotic maps. By launching a low-complexity chosen-plaintext attack, they successfully recovered the equivalent diffusion and permutation keys, thereby completely breaking the cipher and demonstrating that the scheme cannot withstand real-world attacks.

Similarly, Feng et al.[36] performed a systematic cryptanalysis of the IEA-VJD algorithm, which relies on variable-step Josephus traversing and dynamic DNA encoding. They proposed a complete chosen-plaintext attack capable of recovering the plaintext, showing that the algorithm is insecure under realistic attack scenarios. Wen et al.[37] analyzed the CIEA-FOHS image encryption algorithm based on a fractional-order hyperchaotic system and identified several critical flaws, including equivalent keys, reducible permutation, and fragile diffusion. Their work demonstrated that the algorithm can be fully compromised under chosen-plaintext attacks, supported by both theoretical analysis and experimental validation.

Collectively, these studies highlight that many chaotic image encryption schemes—although seemingly robust and exhibiting desirable statistical characteristics—still fail to resist practical threats under rigorous cryptanalytic evaluation. Therefore, newly developed image encryption schemes must not only exhibit strong chaotic dynamical properties but also place emphasis on robustness against various practical attack models.

Although several recent studies have explored the integration of image detection and encryption, most existing works primarily focus on global image processing, while effective fusion of sensitive-target detection and localized encryption remains relatively underexplored.Within a unified security framework, the design of hierarchical data access and differentiated authorization mechanisms is still insufficiently addressed and warrants further investigation.To address these issues, this paper proposes an integrated scheme for target detection and information protection, tailored to the requirements of military aircraft detection and data security in remote sensing scenarios. Built upon the YOLOv7-tiny detection framework and a chaotic encryption architecture, the main innovations and contributions of this work are as follows:

(1) The YOLO-based remote sensing target detection is organically integrated with chaotic encryption, and a dual-layer encryption framework is proposed, consisting of sensitive target encryption followed by global secondary encryption. This design enables a two-tier hierarchical access mechanism for secured data.
(2) A three-branch multi-scale enhanced feature extraction module (MEFABlock) is proposed, which combines multi-scale convolutions and multi-level attention mechanisms to effectively improve the feature representation capability and detection accuracy of the YOLOv7-tiny framework.
(3) A novel chaotic system is designed, and its theoretical chaotic properties and randomness are analyzed, enhancing the complexity of the chaotic sequences and the size of the key space, thereby strengthening the security of the encryption system.
(4) A chaotic initial parameter generation method based on hash scrambling and diffusion is proposed, and an efficient three-level encryption scheme combining improved Josephus scrambling and dual diffusion structure is constructed, further enhancing encryption performance and security reliability.

The remainder of this paper is organized as follows. "Improvements to the detection network based on YOLOv7-tiny" section presents the improved YOLOv7-tiny model. "Chaotic system" section introduces the proposed chaotic system and its corresponding analyses. "Chaotic encryption method" section discusses the chaotic image encryption algorithm. "Experiments and analysis" section reports the experimental results and analysis. "Conclusion" section concludes the paper.

Improvements to the detection network based on YOLOv7-tiny

This section introduces the improved YOLOv7-tiny model, including the design of the MEFABlock and the incorporation of coordinate convolution.

The YOLOv7 model

As a representative one-stage object detection algorithm, the YOLO (You Only Look Once) series has been widely applied in object detection scenarios with high requirements for real-time performance and accuracy, such as remote sensing monitoring and traffic surveillance, owing to its end-to-end design and strong detection performance. The overall architecture of YOLOv7 mainly consists of three parts: Backbone, Neck, and Head[7]. The Backbone extracts low-, middle-, and high-level feature information through multiple convolutional layers; the Neck utilizes multi-scale feature fusion structures (such as FPN and PAN) to efficiently integrate information at different scales, thereby enhancing the detection capability for objects of various sizes; the Head outputs object

locations, categories, and confidence scores through classification and regression branches, realizing end-to-end object detection.

YOLOv7-tiny, as a lightweight version of the YOLOv7 series, further compresses network parameters and reduces computational complexity, making it suitable for deployment on resource-constrained devices and applications with strict inference speed requirements. While maintaining basic detection accuracy, YOLOv7-tiny significantly improves inference efficiency through structural pruning and module simplification strategies, offering excellent practical deployment value.

However, YOLOv7-tiny still exhibits certain limitations in feature representation capability under complex backgrounds, particularly when facing large variations in object scale, dense distributions, and complex interference in remote sensing images. Its detection performance leaves room for further improvement. Moreover, in military aircraft detection and encryption tasks, accurately detecting aircraft targets is a prerequisite for protecting sensitive target information from leakage. Therefore, this paper proposes a Multi-branch Enhanced Feature Aggregation module (MEFABlock) based on the YOLOv7-tiny architecture, aiming to strengthen the network's feature extraction ability and multi-scale information fusion performance. The overall performance of the model for military aircraft detection in remote sensing scenarios is thereby enhanced. The structure of the improved YOLOv7-tiny model is illustrated in Fig. 1.

Multi-branch enhanced feature aggregation block

For aircraft detection tasks in remote sensing imagery, the target datasets present several challenges, including diverse object categories, large variations in structural appearance, significant scale changes, and partially blurred aircraft contours. To address these issues, this paper designs a Multi-branch Enhanced Feature Aggregation Block (MEFABlock) based on the YOLOv7-tiny architecture. The structure of this module is shown in Fig. 2.

The proposed module adopts a parallel convolution structure with 1×1, 3×3, and 5×5 branches to fuse spatial features at different scales, thereby enhancing the network's ability to represent various military aircraft with multi-scale and structural differences. In addition, the lightweight Efficient Channel Attention (ECA) mechanism[38] is introduced to improve the adaptive adjustment of channel-wise feature importance distribution, emphasizing key structural information. Furthermore, the Convolutional Block Attention Module (CBAM)[39], which combines channel and spatial attention, is incorporated to reinforce feature representation from both channel and spatial dimensions, thereby improving feature discrimination and target localization in complex backgrounds.

To balance network representation capability and inference efficiency, the MEFABlock is embedded at the junction between the Backbone and Neck structures in the YOLOv7-tiny architecture. By enhancing deep feature fusion and multi-scale information extraction, the proposed design improves detection performance for military aircraft of different types and scales in complex remote sensing imagery, leading to overall improvements in detection effectiveness and system practicality.

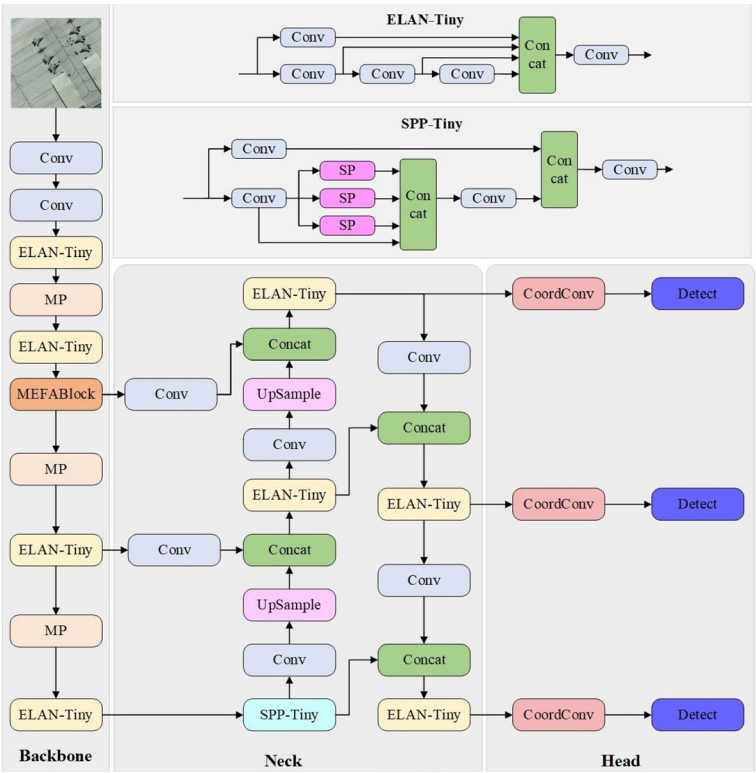


Fig. 1. The structure of improved YOLOv7-tiny.

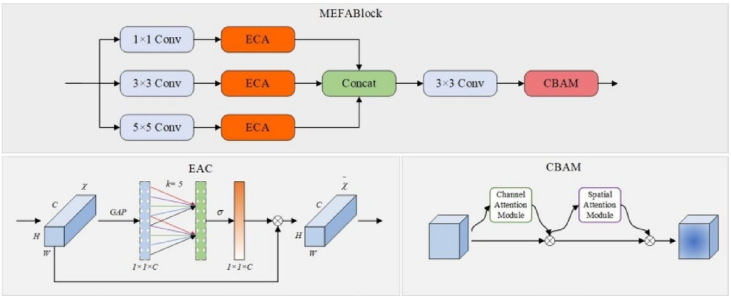


Fig. 2. The structure of MEFABlock.

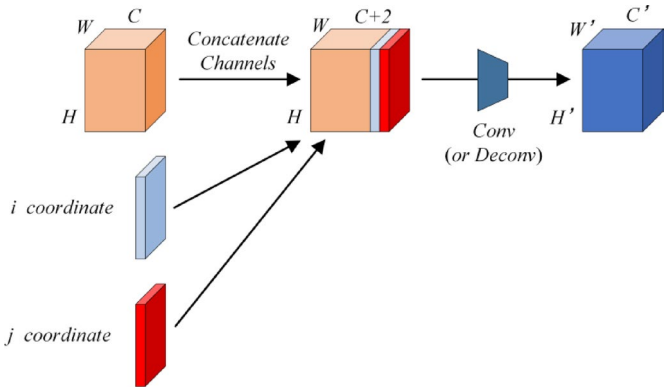


Fig. 3. The structure of CoordConv.

Coordinate convolution

Although YOLOv7-tiny offers good detection speed and basic feature representation capability, conventional convolution operations lack direct modeling of spatial positional information, leading to certain limitations in target boundary localization and spatial structure representation. In aircraft detection tasks in remote sensing imagery, where aircraft types are diverse, contours vary significantly, and scale changes are prominent, network structures without explicit positional awareness still have room for improvement in detection accuracy and localization precision.

To enhance the network's capability to represent spatial positional information, this paper introduces the Coordinate Convolution (CoordConv) structure[40] into the YOLOv7-tiny detection framework, replacing the conventional convolution layers preceding all three YOLOv7-tiny detection heads with CoordConv modules. The structure of CoordConv is illustrated in Fig. 3. CoordConv introduces explicit spatial coordinate information encoded from row and column positions into the feature channels, enabling convolution operations to simultaneously integrate local texture features and global spatial positional information, thereby compensating for the limitations of traditional convolution.

The high-level features before the detection heads contain stronger semantic information and directly affect the final target localization and classification performance. Embedding CoordConv at this position significantly enhances the network's ability to express target positions, scales, and boundary contours while maintaining the overall lightweight architecture. This is particularly beneficial for improving detection accuracy and localization performance for military aircraft of different types and scales in remote sensing scenarios.

Chaotic system

This section presents the proposed 2D-MTCM chaotic system, including its formulation, performance evaluation, and related discussions.

2D-MTCM

The classical one-dimensional Logistic and Sine maps have been widely used in the field of encryption due to their simple structures and basic chaotic properties. However, one-dimensional systems often exhibit periodic windows, have limited key space, and fail to meet the requirements of high-security applications. Therefore, in recent years, researchers have gradually adopted multidimensional coupling and trigonometric functions to enhance the chaotic behavior of systems.

In this paper, inspired by the core principles of Logistic and Sine maps, a novel two-dimensional trigonometric coupled chaotic map, named the 2D Multi-Trigonometric Chaotic Map (2D-MTCM), is proposed, as defined in Eq. (1).

$$\begin{cases} x_{n+1} = [a\sin(\pi x_n) + b\cos(2\pi x_n(1-y_n)) + 0.5(a+b)\sin(y_n)] \bmod 1 \\ y_{n+1} = [b\sin(\pi y_n)\cos(\pi x_{n+1}) + a\cos(\pi y_n(1-x_{n+1})) + 0.5(a+b)\sin(x_n)] \bmod 1. \end{cases} \quad (1)$$

To validate the chaotic properties of the proposed system, this paper conducts experiments including Lyapunov exponent analysis, phase diagram analysis, and 0–1 test, demonstrating that the proposed system meets the requirements for chaotic encryption applications.

Lyapunov exponent analysis

The Lyapunov Exponent (LE) is a key metric used to characterize the sensitivity to initial conditions and the chaotic behavior of a dynamical system. When LE > 0, the system exhibits good chaotic properties[41]. Suppose the chaotic system is represented by $X_{k+1} = F(X_k), X_k = (x_k, y_k)$.Then, the Jacobian can be computed using Eq. (2).

$$J_k = \begin{bmatrix} \frac{\partial f_1}{\partial x}(x_k, y_k) & \frac{\partial f_1}{\partial y}(x_k, y_k) \\ \frac{\partial f_2}{\partial x}(x_k, y_k) & \frac{\partial f_2}{\partial y}(x_k, y_k) \end{bmatrix} . \quad (2)$$

At the k-th iteration, the Jacobian matrix possesses two eigenvalues, denoted by $\lambda_{1,k}$ and $\lambda_{2,k}$.

Accordingly, the Lyapunov exponents of the two-dimensional chaotic system can be evaluated following the formulations given in Eqs. (3)–(4).

$$LE_1 = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln|\lambda_{1,k}|. \quad (3)$$

$$LE_2 = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln|\lambda_{2,k}|. \quad (4)$$

In this study, Lyapunov exponent tests are conducted for the 2D-MTCM system. Since the initial parameters generated by the hash-based initialization method proposed in this paper are guaranteed to be no less than 1, the evaluation of LE_1 is conducted over the parameter range $a, b \geqslant 1$. Based on the evaluation results, the Lyapunov exponent surfaces LE_1 and LE_2 are illustrated in Fig. 4. As shown in Fig. 4, when the initial parameters *a* and *b* are greater than 1, the LE_1 becomes positive, indicating that the system exhibits strong chaotic behavior and is suitable for chaos-based encryption. Moreover, the LE_2 surface reveals that within a certain range of the parameter space, both LE_1 and LE_2 remain positive simultaneously, demonstrating that the proposed system possesses hyperchaotic characteristics.

In addition to the 2D-MTCM system, the chaotic systems proposed in[42–46], as well as the traditional Logistic and Sine chaotic systems, are selected for comparison. The experiment is initialized with specified initial values, and parameters *a* and *b* vary within the range [0.1, 4]. The experimental results are shown in Fig. 5.

As shown in Fig. 5, the LE of the 2D-MTCM system remains greater than 0 and is higher than that of the other systems used for comparison. The 2D-MTCM system consistently maintains a relatively large LE value over a wide parameter range, demonstrating superior chaotic performance and providing a solid theoretical foundation for subsequent encryption applications.

Phase diagram analysis

To verify the ergodicity and chaotic trajectory distribution of the proposed two-dimensional trigonometric coupled chaotic map (2D-MTCM), the phase portraits of the system are plotted under a set of typical parameters,

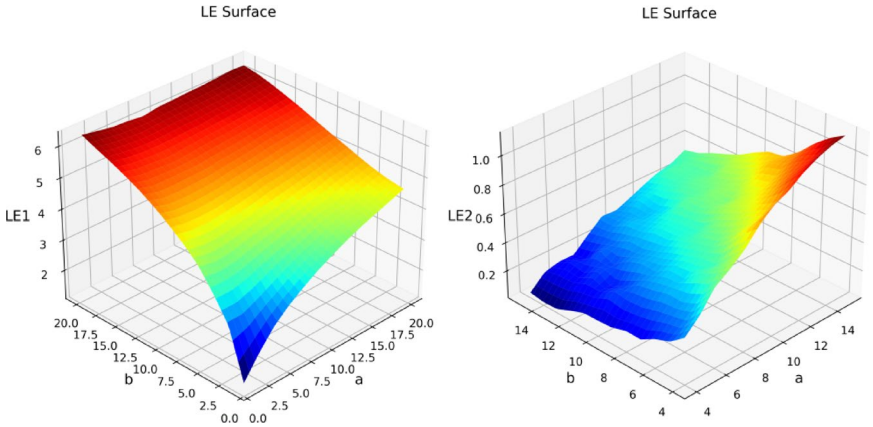


Fig. 4. The LE results.

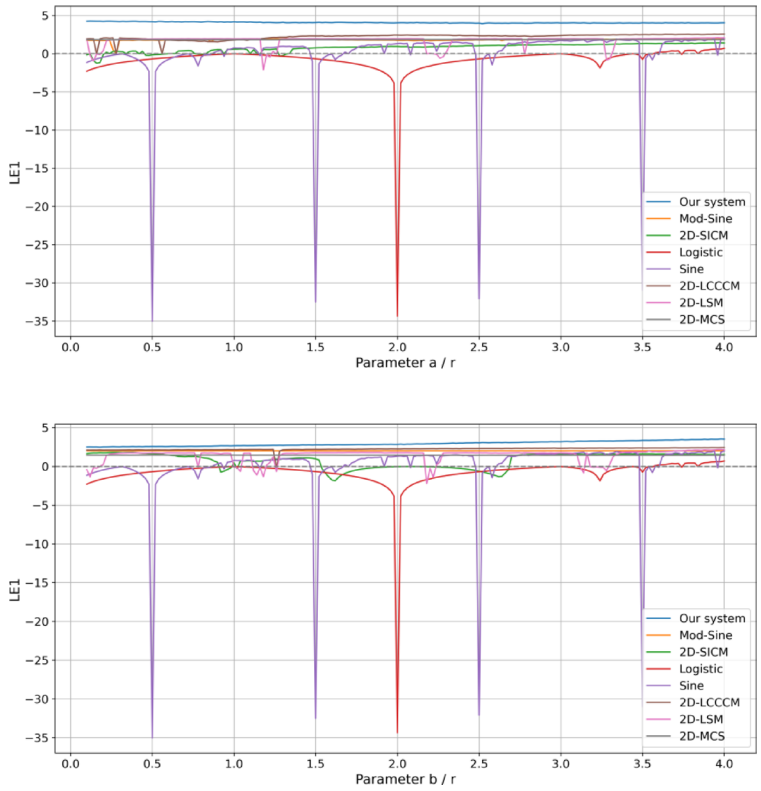


Fig. 5. LE comparison results.

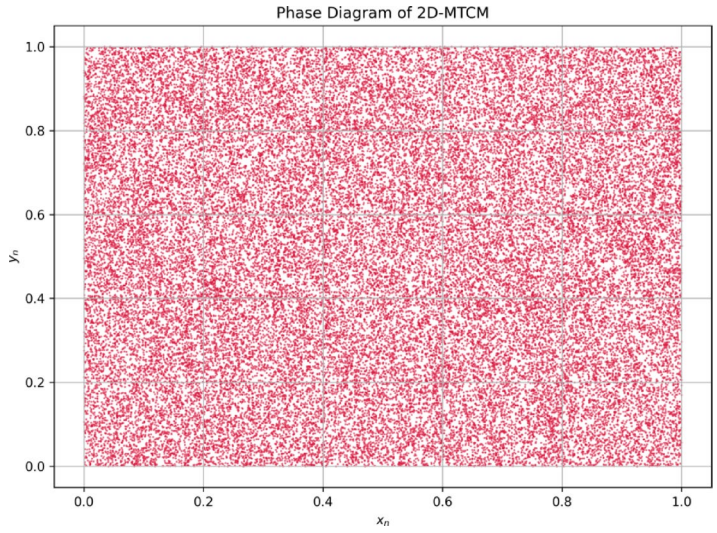


Fig. 6. The phase diagram of 2D-MTCM system.

as shown in Fig. 6. It can be observed that the system state points uniformly cover the entire phase space, exhibiting good ergodicity and chaotic characteristics.

Bifurcation diagram analysis

To further analyze the dynamical characteristics of the proposed two-dimensional trigonometric coupled chaotic map (2D-MTCM) under parameter variations, bifurcation diagrams of x_n with respect to parameters a and b are plotted, as shown in Fig. 7. It can be observed that the system exhibits strong chaotic behavior and good ergodicity across a wide range of parameter values, with trajectories densely distributed in the interval [0, 1], and no obvious periodic windows or stable attractors appearing. Compared with traditional chaotic systems,

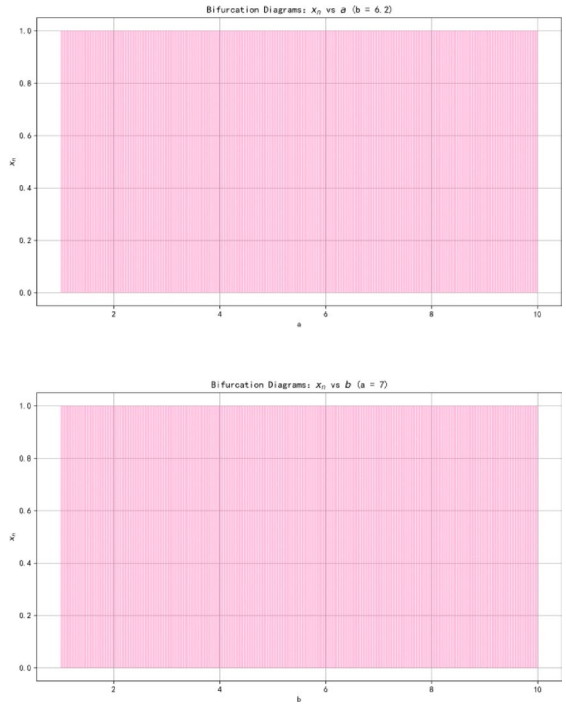


Fig. 7. Bifurcation Diagrams of the 2D-MTCM System.

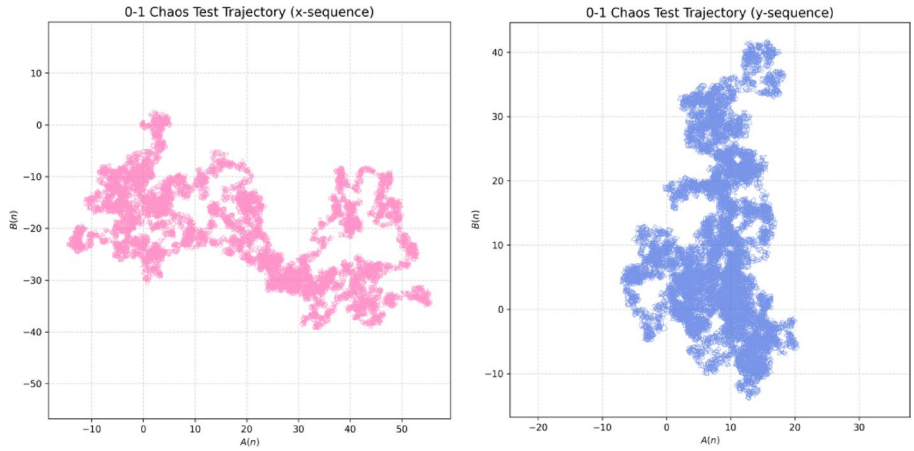


Fig. 8. 0–1 Test results of the 2D-MTCM system.

the proposed system demonstrates higher unpredictability and superior parameter adaptability, providing a sufficiently large key space and enhanced security assurance for encryption applications.

0–1Test

To further verify the chaotic behavior of the proposed two-dimensional trigonometric coupled chaotic map (2D-MTCM), the classical 0–1 test for chaos is employed. The 0–1 tests are performed separately on the chaotic sequences of the x-component and y-component, and their diffusion trajectories are shown in Fig. 8. For the 0–1 test, the definitions of the variables are provided in Eqs. (5)–(8).

Given the sequence x_n,the phase is defined as shown in Eq. (3),where $r = \frac{\pi}{4}$.

$$\theta(n) = r \cdot n + \sum_{j=1}^{n} x_j. \tag{5}$$

Subsequently, two cumulative sums, A(n) and B(n), are constructed.

$$A(n) = \sum_{j=1}^{n} x_j \cos(\theta(j)). \tag{6}$$

$$B(n) = \sum_{j=1}^{n} x_j \sin(\theta(j)). \tag{7}$$

The mean square displacement M(n) is then defined.

$$M(n) = A(n)^2 + B(n)^2. \tag{8}$$

The results show that the trajectories of (A(n), B(n)) exhibit obvious diffusion-like random walk characteristics in phase space, without periodicity or convergence phenomena. This Brownian motion-like trajectory pattern indicates that the system possesses good chaotic diffusion properties.

In addition, the 0–1 test K-values for the *x*-component and *y*-component chaotic sequences are calculated to be K = 0.8996 and K = 0.7613, respectively, with the corresponding linear slopes of M(n) being 0.0644 and 0.0210. A K-value closer to 1 indicates stronger chaotic behavior. The experimental results demonstrate that the chaotic sequences generated by the proposed system exhibit high unpredictability and complexity, meeting the requirements of encryption algorithms for high chaos and security of keystreams.

NIST SP800-22 test

To further validate the statistical properties of the keystreams generated by the proposed two-dimensional trigonometric coupled chaotic map (2D-MTCM), NIST SP800-22 standard randomness tests were performed on 20 million bits of keystreams from both the *x*-component and *y*-component. As shown in Table 1, all main test items and their sub-tests yielded *P*-values significantly greater than 0.01, and the pass rates met or exceeded the minimum requirements specified by NIST.

Both the *x*-component and *y*-component demonstrated excellent performance in all statistical tests, indicating that the proposed system can reliably generate high-strength, highly uniform pseudorandom keystreams, fully meeting the randomness and security requirements for practical encryption applications.

Chaotic encryption method

This section provides a detailed description of the proposed chaotic encryption algorithm, including its implementation procedure and technical details.

Encryption workflow

To effectively protect key information in images, this paper proposes a chaotic encryption algorithm based on a diffusion–permutation–diffusion structure. The overall workflow is shown in Fig. 9. The algorithm employs the two-dimensional trigonometric coupled chaotic system (2D-MTCM) as its core and significantly enhances the security and attack resistance of the ciphertext through a three-stage progressive process.

First, the algorithm applies SHA-512 and SHA-256 hash functions to the original image and its individual channel information to dynamically generate the initial parameters and control variables of the chaotic system, thereby achieving an "image-dependent + key-dependent" chaotic sequence driving mechanism. This

No.	Test item	X-sequence *P*-value	X pass rate	Result	Y-sequence *P*-value	Y pass rate	Result
1	Frequency	0.276	0.95	Pass	0.911	1.00	Pass
2	Block frequency	0.122	1.00	Pass	0.213	1.00	Pass
3	Cumulative sums F/B	0.122/0.213	0.95	Pass	0.637/0.834	1.00	Pass
4	Runs	0.534	1.00	Pass	0.911	0.95	Pass
5	Longest run	0.350	1.00	Pass	0.534	1.00	Pass
6	Rank	0.350	0.95	Pass	0.049	1.00	Pass
7	FFT	0.964	1.00	Pass	0.276	1.00	Pass
8	Non-overlapping template	0.421	1.00	Pass	0.421	1.00	Pass
9	Overlapping template	0.163	0.95	Pass	0.638	0.95	Pass
10	Universal	0.437	1.00	Pass	0.834	0.95	Pass
11	Approximate entropy	0.911	1.00	Pass	0.049	1.00	Pass
12	Serial	0.534/0.834	1.00	Pass	0.834/0.637	1.00	Pass
13	Linear complexity	0.534	1.00	Pass	0.911	1.00	Pass
14	Random excursions	0.521	1.00	Pass	0.521	1.00	Pass
15	Random excursions variant	0.439	1.00	Pass	0.439	1.00	Pass

Table 1. NIST SP800-22 test results. The *P*-values represent typical or average values, and all sub-test pass rates meet the NIST standards.

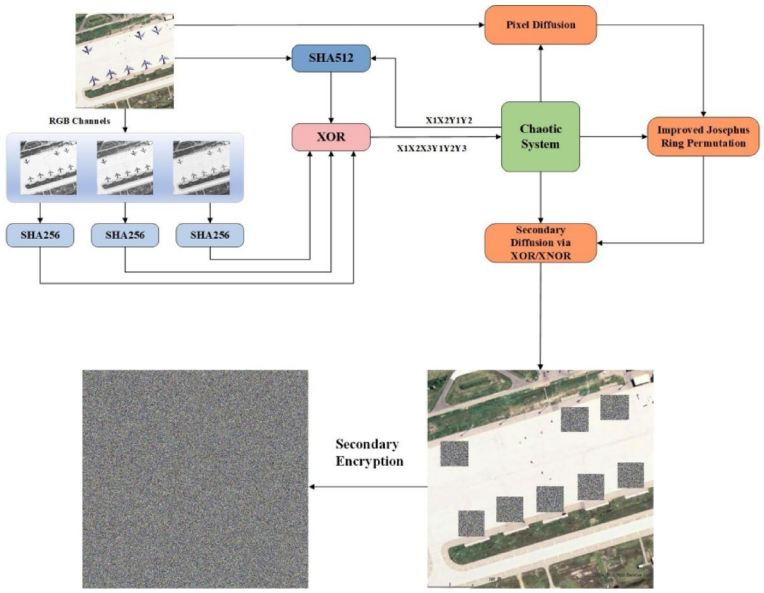


Fig. 9. Overall encryption workflow.

mechanism significantly increases the key space and unpredictability of the algorithm, enhancing its resistance to attacks.

Next, the encryption process is carried out in three steps: The first diffusion stage employs chaotic sequences to globally perturb pixel values, achieving high diffusion of pixel intensities.The second stage adopts the Josephus ring scrambling strategy to permute pixel positions, disrupting spatial correlations.The third stage performs a second diffusion operation, further increasing the complexity of the correlation between plaintext and ciphertext.

This encryption framework not only effectively prevents conventional statistical and differential analysis attacks but also exhibits good parallelism and scalability, making it suitable for protecting various types of sensitive image data.

Hash-based chaotic parameter generation method

First, the algorithm performs SHA-512 and SHA-256 hash operations on the original image data and its individual channel information to obtain high-strength, collision-resistant hash digests. Then, the obtained hash values are segmented and concatenated using a fixed-length window, and mixed through bitwise XOR and other operations to generate multiple parameter seeds K1 to K8.

These K1–K8 parameter seeds are used to generate the initial parameters for chaotic sequences X10, Y10, X20, and Y20. After generating two groups of chaotic sequences using the chaotic system, iterative operations are performed on these sequences, and the sum of the first 3,000 digits of X10, Y10, X20, and Y20 is computed to produce a new SHA-512 hash.

This new SHA-512 hash is then processed in the same manner to obtain another set of parameter seeds K1–K8. These second-round seeds are normalized and used to generate the initial values and control parameters for the two-dimensional chaotic system that participates in the encryption process.

Taking X10 and Y10 as examples, the initial parameter generation formulas are given in Eqs. (9)–(12).

$$X10 = \frac{SumHex(K_1 \oplus K_2 \oplus K_3) + SumHex(K_4)}{1024}. \tag{9}$$

$$Y10 = \frac{SumHex(K_5 \oplus K_6 \oplus K_7) + SumHex(K_8)}{1024}. \tag{10}$$

$$a_{10} = \max\left(\frac{SumHex(K_1 \oplus K_2 \oplus K_3) + SumHex(K_4 \oplus K_5)}{256}, 1\right). \tag{11}$$

$$b_{10} = \max\left(\frac{SumHex(K_5 \oplus K_6 \oplus K_7) + SumHex(K_8 \oplus K_1)}{256}, 1\right). \tag{12}$$

Here, *SumHex(·)* denotes the operation of converting each hexadecimal character of the hash string into decimal and summing them, and$\oplus$represents the bitwise XOR operation. For generating the initial parameters of subsequent chaotic sequences, taking Eq. (9) as an example, for X20, the K seed group is circularly left-shifted by one position, i.e., X20 is generated using K2, K3, K4, and K5, and so on for the others.The pseudocode for this step is shown in Table 2.

Input：RGB image I of size H×W
Output：Three sets of chaotic parameters
1: Hashing:
2: Compute H_all (SHA-512 of I), H_R, H_G, H_B (SHA-256 of each channel).
3: Mixing:
4: Form 128-hex mixed hash M by segment-wise XOR of H_all, H_R, H_G, H_B.
5: Grouping:
6: Divide M into 8 groups (K_1~K_8), each with 32 hex digits.
7: Initial Parameter Calculation:
8: Compute (a, b, x_0, y_0) from K_1~K_8 via:
9: - x_10 = (SumHex($K_1 \oplus K_2 \oplus K_3$) + SumHex($K_4$)) / 1024
10: - y_10 = (SumHex($K_5 \oplus K_6 \oplus K_7$) + SumHex($K_8$)) / 1024
11: - a_{10} = max((SumHex($K_1 \oplus K_2 \oplus K_3$) + SumHex($K_4 \oplus K_5$)) / 256, 1)
12: - b_{10} = max((SumHex($K_5 \oplus K_6 \oplus K_7$) + SumHex($K_8 \oplus K_1$)) / 256, 1)
13: Hash Cascading:
14: Generate a chaotic sequence of length N+3000, hash its first 3000 values to obtain a new 128-hex hash.
15: Final Parameters:
16: Repeat steps 3–4 using the new hash, with left shifts (shifts 1, 2, 3) to generate three sets of (a_{10}, b_{10}, x_10, y_10) for encryption.

Table 2. Pseudocode for generating chaotic system initial parameters from hash values.

As a result, three groups of chaotic sequences, X1, Y1, X2, Y2, X3, and Y3, are finally obtained and are used to guide the encryption operations for the first diffusion, permutation, and second diffusion, respectively.

Pixel diffusion

In the first diffusion process, two chaotic sequences X1 and Y1, each with a length equal to the number of pixels, are used to perturb the pixel values. The specific steps are as follows:

First, for each pixel position *i*, the chaotic perturbation amounts are calculated as shown in Eqs. (13) and (14).

$$X\_\mathrm{mod} = ((x[i] + y[i]) \times 10^{18}) \bmod 256. \tag{13}$$

$$Y\_\mathrm{mod} = ((x[i] - y[i]) \times 10^{18}) \bmod 256. \tag{14}$$

Then, the original pixel value is XORed with the two perturbation values to obtain the intermediate diffusion result, as shown in Eq. (15).

$$N = Original\ Pixel[i]\ XOR\ X\_\mathrm{mod}\ \ XOR\ Y\_\ \mathrm{mod}. \tag{15}$$

To further enhance the nonlinearity and security of the diffusion process, an additional perturbation is introduced using Eq. (16).

$$Z = (\frac{(x[i] + y[i])}{(|x[i] - y[i]| + \varepsilon)} \times 10^{18}) \bmod 256. \tag{16}$$

Here, ε is set as a very small positive value to avoid division by zero. The final diffusion output is defined as shown in Eq. (17).

$$E = (N + Z) \bmod 256. \tag{17}$$

Improved Josephus ring permutation

The Josephus ring was originally a mathematical problem, and applying the Josephus ring to chaotic image encryption can enhance the resistance of the encryption scheme against brute-force attacks[47]. In the second step of the proposed method, an improved Josephus ring permutation mechanism driven by chaotic sequences is introduced to achieve global rearrangement of pixel positions.

The specific procedure is as follows: First, the input channel is flattened into a one-dimensional array with a total length of *N*. A position index queue of length *N* is then constructed, initialized as [0,1,2,…,*N*−1], to simulate a circular structure. In this scrambling process, at each step, a jump length is generated from the chaotic sequence X2. Starting from the current position, the index queue is traversed circularly to select a position, and the pixel at the selected position is added to the output sequence, while the corresponding element is removed from the queue. This procedure is repeated until all pixels are selected. The scrambled pixel sequence is then reshaped back to the original channel dimensions. The computation of the jump length *J* is given in Eq. (18), where M = max(rows, columns), M′=ceil[M/2], and M′′=floor[M/2]. Here, Ceil denotes the ceiling operation, and floor denotes the floor operation. After every M selections, Eq. (18) is recalculated to obtain a new *J*. To avoid overflow, 1 is subtracted from *J* after each update. If *J* = 0, it is reassigned to 1 to prevent infinite loops; if a negative value occurs, the absolute value of *J* is taken before proceeding.

$$J = floor(chaotic\_seq[i] \times 10^{16} + \frac{chaotic\_seq[i]}{2} \times 10^{18}) \bmod\ \ M'' + M'. \tag{18}$$

Second pixel diffusion

To further enhance the security of the encrypted image and increase the strength of pixel perturbation, a second diffusion process based on chaotic sequences and XOR/XNOR operations is applied after the first diffusion and permutation. This process effectively breaks the correlations between pixels and enhances the statistical properties of the ciphertext.

Let the input be the permuted single-channel image C of size H × W, with a total of N pixels. The second pixel diffusion uses chaotic sequences X3 and Y3, whose elements are denoted as x[i] and y[i], respectively. The operations are performed as defined in Eqs. (19) and (20).

$$A_i = \left(\frac{(x[i] + y[i])}{2} \times 10^{10} \right) \bmod 256. \tag{19}$$

$$B_i = \left(\sqrt{x[i] + y[i]} \times 10^{10} \right) \bmod 256. \tag{20}$$

For each pixel Ci (flattened into a one-dimensional sequence), if A[i] ≥ B[i], the operation defined in Eq. (21) is performed; otherwise, the operation defined in Eq. (22) is executed. In Eqs. (21)-(22), $\oplus$represents the bitwise XOR operation,$\odot$represents the bitwise XNOR operation.

$$E_i = C_i \oplus (A_i \odot B_i). \tag{21}$$

$$E_i = (C_i \odot A_i) \oplus B_i. \tag{22}$$

All E[i] values are then reshaped into an H × W two-dimensional matrix to obtain the image after the second diffusion. The above process is performed independently for the R, G, and B channels, and the final encrypted image is obtained by merging the three encrypted channels.

Through this three-step diffusion–permutation–diffusion process, the image is encrypted. The decryption process involves inputting the encrypted ciphertext image, reading the hash-based key, and performing the inverse operations of the three steps described above to ultimately recover the decrypted image. During decryption, the hash key is input, and the chaotic system generates the same chaotic sequence based on the key. The process then applies inverse second diffusion, inverse permutation, and inverse first diffusion sequentially, thereby restoring the encrypted pixels to the original image pixels.

Experiments and analysis

This section presents the experiments and analyses of the proposed method, including the evaluation of the improved YOLOv7-tiny model and the performance tests of the chaotic encryption algorithm.

Dataset and experimental settings

The dataset used in this study is the MAR20 military aircraft recognition dataset developed by Northwestern Polytechnical University[48]. This dataset contains a total of 3842 images and is divided into training and testing subsets. The training set consists of 1,331 images and 7870 target instances, while the testing set contains 2,511 images and 14,471 target instances.

The MAR20 dataset includes 20 categories of military aircraft models, which are labeled A1 through A20 by the authors. Example images from the dataset are shown in Fig. 10, and the distribution of categories in the MAR20 dataset is illustrated in Fig. 11.

The experimental environment and configurations used in this study are summarized in Table 3.

In the experiments of this study, the main hyperparameter settings for training the YOLOv7-tiny model are as follows: the initial learning rate (lr0) is set to 0.01, the final learning rate factor (lrf) is 0.1, the momentum is 0.937, and the weight decay is 0.0005.

Fig. 10. Example image from the MAR20 dataset.

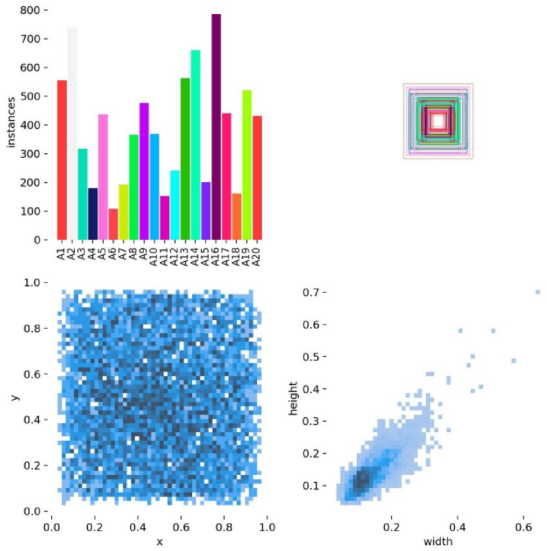


Fig. 11. Category distribution of the MAR20 dataset.

Item	Specification
Operating system	Win10
CPU	Intel(R)Core(TM)i5-14600KF
RAM	32GB
GPU	NVIDIA GeForce RTX 4080super
CUDA	12.1
Python	3.9
Pytorch	2.2.2
Epoch	400
Batch-size	16
optimizer	SGD

Table 3. Experimental configurations in this study.

Model comparison experiments

Precision (P), recall (R), F1-score, and mean average precision (mAP) are selected as evaluation metrics in this study. The specific calculation formulas are provided in Eqs. (23)–(27).

$$P = \frac{TP}{TP + FP} \times 100\%. \tag{23}$$

$$R = \frac{TP}{TP + FN} \times 100\%. \tag{24}$$

$$F1 = \frac{2 \times (P \times R)}{P + R}. \tag{25}$$

$$AP = \int_0^1 P(r)dr. \tag{26}$$

$$mAP = \frac{\sum_{i-1}^{C} AP_i}{C}. \tag{27}$$

In the given equations, TP (true positives) denotes the number of correctly identified positive instances, while FP (false positives) represents cases where negative samples were mistakenly classified as positive. FN (false negatives) accounts for the positive instances that the model failed to identify. The average precision (AP) measures the area under the precision-recall curve for a specific class, and the mean average precision (mAP) summarizes the overall detection accuracy by averaging AP values across all categories.

Algorithm	*P*/%	*R*/%	F1-score	mAP@0.5	mAP@0.5:0.95
RT-DETR[49]	83.1	78.5	80.7	77.7	58.7
YOLOv5n[5]	81.6	78.0	79.8	85.3	64.4
YOLOv6n[6]	77.9	74.9	76.4	80.9	60.8
YOLOv7-tiny[7]	83.6	80.2	81.9	87.3	64.4
YOLOv8n[8]	82.5	78.9	80.7	85.7	64.2
YOLOv9t[9]	77.6	75.6	76.6	82.6	62.3
YOLOv10n[10]	82.8	78.5	80.6	85.9	65.2
YOLOv11n[11]	82.1	79.0	80.5	86.1	64.7
YOLOv12n[12]	79.9	78.5	79.2	85.0	64.0
YOLOv13n[13]	77.9	73.3	75.5	80.7	60.6
Ours	**86.0**	**82.9**	**84.4**	**90.0**	**66.6**

Table 4. Comparison of different object detection models. Significance value bold.

	RE-DETR	v5n	v6n	v7-tiny	v8n	v9t	v10n	v11n	v12n	v13n	Ours
ALL	77.7	85.3	80.9	87.3	85.7	82.6	85.9	86.1	85.0	80.7	**90.0**
A1	73.9	82.4	76.1	80.7	78.4	75.4	86.7	81.4	74.1	72.9	**88.0**
A2	81.3	93.6	91.8	97.1	94.5	90.7	94.0	93.9	93.2	90.8	**98.2**
A3	86.9	89.8	89.0	95.9	92.9	89.7	93.2	90.0	93.2	88.0	**98.2**
A4	82.6	88.5	87.0	93.1	91.6	87.9	88.6	90.2	91.1	86.8	**95.6**
A5	69.8	70.2	66.3	68.8	76.4	67.2	71.4	72.6	73.5	65.8	**75.1**
A6	91.0	96.4	92.7	96.7	96.6	94.5	95.9	**97.5**	95.5	89.5	96.0
A7	90.9	95.9	90.3	95.4	96.2	94.7	94.2	96.4	**97.2**	94.9	96.5
A8	87.1	**94.7**	92.8	90.0	92.8	94.3	94.1	93.0	**94.7**	93.4	94.5
A9	76.1	90.0	81.7	91.7	91.6	83.7	87.2	88.8	88.0	80.9	**94.8**
A10	95.0	97.0	95.4	98.2	97.6	94.9	96.9	97.4	97.6	95.2	**98.8**
A11	70.8	83.2	75.1	81.6	78.1	79.0	84.7	**86.8**	83.3	81.8	83.3
A12	75.1	89.0	81.5	90.3	88.0	83.4	89.0	90.0	82.5	74.9	**93.2**
A13	62.5	76.5	76.4	80.8	79.2	74.3	80.2	79.9	80.6	74.8	**82.7**
A14	82.9	93.3	90.3	92.9	93.3	92.0	93.3	92.7	92.0	90.9	**94.1**
A15	39.3	45.7	28.3	58.0	47.7	48.8	45.3	50.6	**63.5**	49.5	57.6
A16	76.9	92.1	**94.1**	**94.1**	93.6	88.1	93.2	**94.1**	91.7	88.7	93.8
A17	88.1	93.9	94.0	97.1	96.3	92.4	95.2	95.9	93.2	90.7	**98.0**
A18	72.2	77.4	60.4	82.4	76.1	71.6	77.4	71.9	69.8	62.0	**87.0**
A19	75.0	77.0	75.4	78.4	76.5	73.9	77.6	79.8	72.7	68.6	**85.5**
A20	77.0	79.0	79.4	83.6	77.1	76.0	80.6	79.5	72.1	74.1	**89.6**

Table 5. Class-wise performance comparison of different models on subcategories A1–A20 and the overall average (ALL). Significance value bold.

The proposed improved algorithm is compared with other mainstream YOLO series algorithms, and the experimental results are shown in Table 4. From the results in Table 4, it can be observed that the improved algorithm achieves the best performance in terms of precision (P), recall (R), F1-score, and mean average precision (mAP@0.5 and mAP@0.5:0.95).

As shown in Table 4, in addition to comparing various YOLO series algorithms (YOLOv5n to YOLOv13n), this study also introduces RT-DETR[49], a representative detection model in recent years, as well as the latest YOLOv13 model, forming a comprehensive performance comparison framework covering multiple versions and structural types.

Compared to the baseline YOLOv7-tiny model, the proposed improved model achieves significant improvements in all evaluation metrics: precision (P) increases from 83.6% to 86.0%, recall (R) increases from 80.2% to 82.9%, F1-score improves to 84.4%, mAP@0.5 improves to 90.0%, and mAP@0.5:0.95 reaches 66.6%, ranking first among all evaluated models. In addition, we evaluated the real-time capability of the improved network. With only 6.7 M parameters and 21.2 GFLOPs, the model achieves an end-to-end speed of 149 FPS, meeting the real-time demands.

Table 5 presents the comparison of different algorithms on the 20 military aircraft categories, with mAP@0.5 as the primary metric to assess overall detection accuracy. The bold numbers in the table indicate the best-performing algorithm for each category.

To more intuitively illustrate the differences in detection accuracy for the 20 categories of military aircraft, a radar chart is plotted as shown in Fig. 12. Combined with Table 5; Fig. 12, it can be seen that the proposed algorithm achieves the highest accuracy in identifying 14 categories of military aircraft.

YOLOv7-tiny still has room for improvement in detection accuracy due to its lack of efficient multi-scale feature fusion and spatial perception capabilities. In contrast, YOLOv8n–YOLOv13n show continuous evolution in feature modeling and architectural optimization but still fail to outperform the proposed model on the MAR20 military aircraft dataset.

Although RT-DETR possesses end-to-end detection capabilities and strong global modeling ability, it suffers from insufficient boundary clarity and inadequate spatial detail capture in remote sensing tasks, resulting in significantly lower mAP@0.5:0.95 compared to the YOLO series.

By introducing the MEFABlock module to enhance feature representation and incorporating CoordConv before the detection heads to improve spatial information modeling, the proposed model achieves a significant performance improvement in military aircraft detection tasks under remote sensing backgrounds, demonstrating the effectiveness and adaptability of the architectural design.

To further illustrate the effectiveness of the improved YOLOv7-tiny in military aircraft detection, comparative experiments on military aircraft target detection in remote sensing images were conducted, with results shown in Table 6.

As seen from Table 6, in actual comparison image (a), structurally advanced models such as YOLOv11n and YOLOv13n failed to detect the aircraft, whereas the proposed improved model successfully detected it with correct classification. In actual comparison image (b), the YOLOv11n model exhibited a missed detection, and both YOLOv11n and YOLOv13n incorrectly identified the aircraft as category A20, while according to manually annotated ground truth, the true category was A13; the improved model successfully detected and classified it correctly. In actual comparison image (c), all models performed relatively well, but the improved model achieved higher overall detection confidence than YOLOv11n and YOLOv13n.

Ablation experiments

To investigate the contribution of the proposed modules to the improvement in model accuracy, ablation experiments were conducted. The results of the ablation experiments are shown in Table 7. In Table 7, Module A represents MEFABlock, Module B represents coordinate convolution (CoordConv), and "✓" indicates that the corresponding module is included in the experimental setting.

As shown in Table 7, the baseline YOLOv7-tiny model, without any structural improvements, achieves an mAP@0.5 of 87.3% and an mAP@0.5:0.95 of 64.4%. When only MEFABlock is introduced, precision increases to 85.9%, mAP@0.5 improves to 89.3%, and mAP@0.5:0.95 reaches 66.1%, indicating that the incorporation of the three-branch convolution fusion structure and attention mechanisms significantly enhances multi-scale feature extraction capability.

When only coordinate convolution is introduced, recall increases from 80.2% to 82.5%, demonstrating that explicit positional information improves the model's ability to capture aircraft edge contours, which is particularly suitable for the large scale variations among different types of aircraft in the MAR20 dataset. mAP@0.5 improves to 88.7% and mAP@0.5:0.95 rises to 65.5%, also showing positive gains.

When both modules are used together, the model achieves the best performance across all metrics: precision increases to 86.0%, recall increases to 82.9%, mAP@0.5 reaches 90.0%, and mAP@0.5:0.95 improves to 66.6%. These results indicate that the proposed feature enhancement module and spatial perception mechanism exhibit good structural complementarity, collaboratively improving target detection accuracy and boundary localization in military aircraft remote sensing images.

To further investigate the contributions of each component within the MEFABlock, ablation experiments were conducted on its individual structures. The experimental results are presented in Table 8.

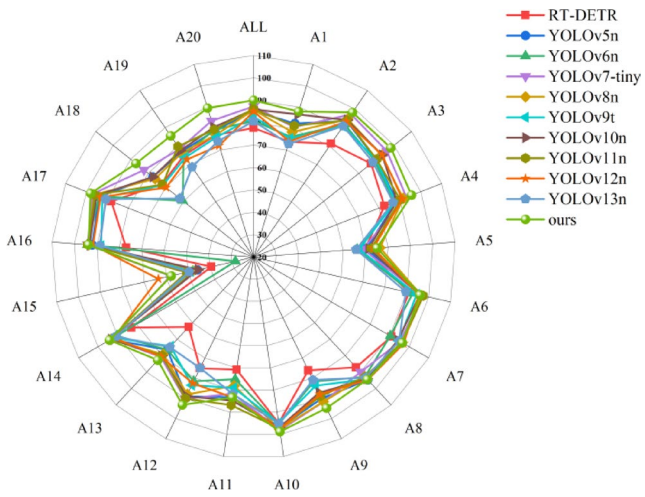


Fig. 12. Radar chart comparing model accuracy on the MAR20 dataset.

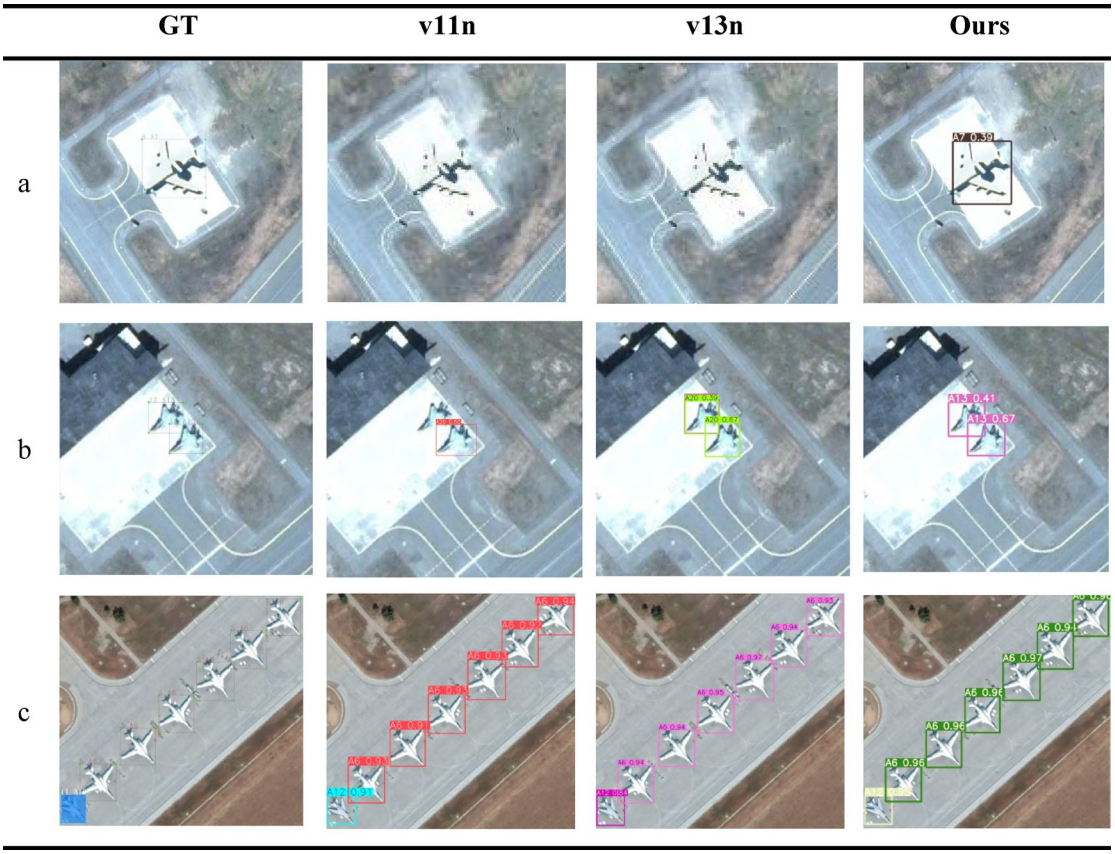


Table 6. Actual detection comparison on remote sensing images.

Baseline	A	B	*P*/%	*R*/%	F1-score	mAP@0.5	mAP@0.5:0.95
✓			83.6	80.2	81.9	87.3	64.4
✓	✓		85.9	82.8	84.3	89.3	66.1
✓		✓	84.1	82.5	83.3	88.7	65.5
✓	✓	✓	**86.0**	**82.9**	**84.4**	**90.0**	**66.6**

Table 7. Ablation experiments. Significance value bold.

Baseline	Three-branch	+ECA	+CBAM	*P*/%	*R*/%	F1-score	mAP@0.5	mAP@0.5:0.95
✓				83.6	80.2	81.8	87.3	64.4
✓	✓			85.1	82.8	83.9	89.1	65.7
✓	✓	✓		85.3	83.2	84.2	89.4	65.9
✓	✓		✓	86.0	82.5	84.2	89.3	66.0
✓	✓	✓	✓	85.9	82.8	84.3	89.3	66.1

Table 8. MEFABlock ablation experiments. Significance value bold.

As shown in Table 8, adding the ECA attention mechanism on top of the three-branch structure leads to improvements in both Recall (R) and mAP, indicating that ECA enhances the response of key semantic feature channels and provides a mild yet positive optimization effect on the model. When introducing the CBAM module alone, the recall decreases slightly by 0.3% compared with the standalone three-branch structure; however, the Precision (P) exhibits a substantial increase. This suggests that CBAM effectively selects informative semantic channels and suppresses background noise in the spatial domain, thereby highlighting target regions. When both ECA and CBAM attention mechanisms are incorporated simultaneously, the model achieves the highest F1 score, attaining an optimal balance between P and R. Furthermore, this configuration yields the best

performance in mAP@0.5:0.95, demonstrating that the combination of the three-branch structure with both attention mechanisms offers the most favorable overall performance.

The ablation experiment results fully validate the independent and combined effectiveness of MEFABlock and CoordConv in improving detection performance, and demonstrate the rationality of the architectural design proposed in this paper.

Chaotic encryption method tests

Key space analysis and key sensitivity analysis

The key used in this study is derived from SHA-512 computation, with a key length of 2^{1280} bits, which exceeds the 2^{100} requirement[50]. Therefore, the proposed algorithm possesses a large key space, enabling it to resist brute-force attacks.

A key sensitivity analysis of the encryption algorithm was performed as follows:

The original key A is:

5ee3a198052db72355dc174cd56cb6297e9cb85c02897d936cb89b77

c78cbd7eee9cf4166b43f951576e32ac2393390e1907229d430305ea7e147ed74aa1d23f.

By changing the first bit, key B is obtained:6ee3a198052db72355dc174cd56cb6297e9cb85c02897d936cb89b77c78cbd7eee9cf4166b43f951576e32ac2393390e1907229d430305ea7e147ed74aa1d23f.

By changing the last bit, key C is obtained:5ee3a198052db72355dc174cd56cb6297e9cb85c02897d936cb89b77c78cbd7eee9cf4166b43f951576e32ac2393390e1907229d430305ea7e147ed74aa1d23e.

During decryption, the original key, key B, and key C were used respectively to attempt decryption. The results are shown in Table 9.

As shown in Table 9, although the key was altered only slightly (by a single bit), the decrypted images remained ciphertext. Therefore, the proposed encryption algorithm exhibits good key sensitivity.

Histogram analysis

To evaluate the proposed encryption algorithm's ability to obfuscate the pixel value distribution of images, histogram analysis was performed on both the original and encrypted images. Histograms can intuitively reflect the distribution patterns of pixel intensity values in an image and are one of the commonly used methods for evaluating encryption effectiveness.

The pixel value distributions were computed for each channel (R, G, and B) of both the original and encrypted images. For each channel, the frequency of pixel values in the range 0–255 was counted and the corresponding histograms were plotted. The histogram results are shown in Table 10.

As shown in Table 10, the histogram of the original image exhibits obvious statistical patterns, with a large number of pixels concentrated at specific gray levels, reflecting the structural characteristics of the original image content. In contrast, the histogram of the encrypted image shows a nearly uniform distribution, with pixel values evenly spread over the range 0–255 and almost equal frequency for all gray levels, indicating that the encryption process effectively disrupts the statistical properties of the original image.

The histogram analysis results demonstrate that the proposed encryption algorithm can significantly enhance the uniformity of pixel values and effectively conceal the statistical information of the original image.

Information entropy analysis

Information entropy is an important metric for measuring the randomness and uncertainty of image information, with a theoretical maximum value of 8. For encrypted images, the closer the information entropy is to this theoretical maximum, the more random the pixel distribution, the higher the security of the encryption algorithm, and the stronger its resistance to statistical attacks. The formula for calculating information entropy is given in Eq. (28). In Eq. (28), $P(i)$ denotes the probability that the pixel takes the gray level.

In this study, five images were selected for information entropy testing: Baboon (512 × 512), Pepper (512 × 512), Airplane (512 × 512), and Remote sensing image1(RSimage1,825 × 799), Remote sensing image2(RSimage2800 × 800).The RSimage1 and RSimage2 are the actual remote sensing military aircraft detection image used in this paper. The results of the information entropy tests are presented in Table 11.

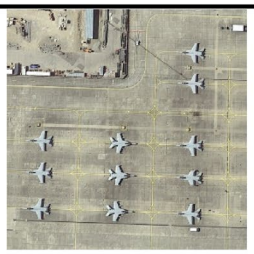
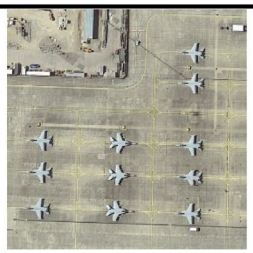
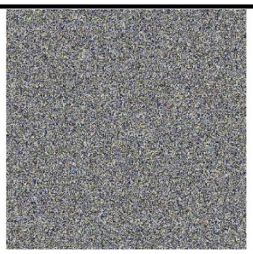
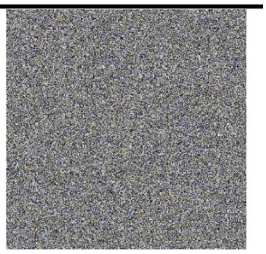
Original image	Key A	Key B	Key C
			

Table 9. Experimental results of key sensitivity analysis.

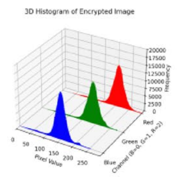
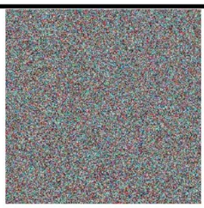
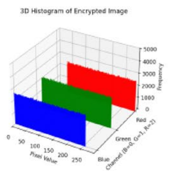
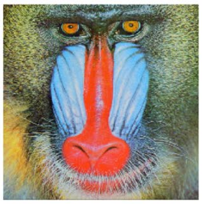
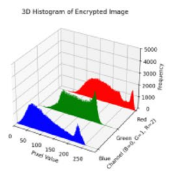
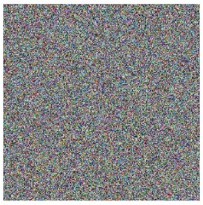
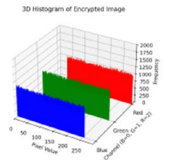
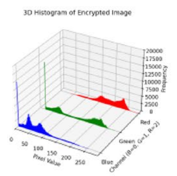
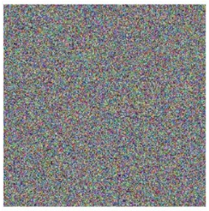
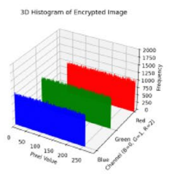
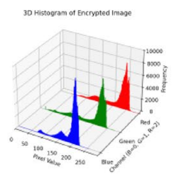
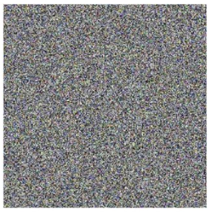
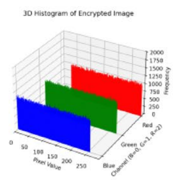
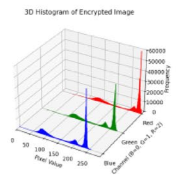
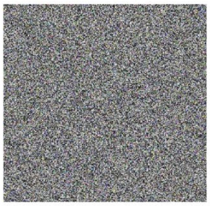
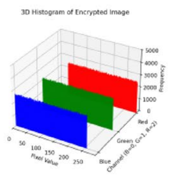
Original image	Histogram of original image	Encrypted image	Histogram of encrypted image
			
			
			
			
			

Table 10. Histogram analysis Results.

Image	*R*	G	B	Mean
Baboon	7.9992	7.9992	7.9994	7.9993
Pepper	7.9993	7.9993	7.9994	7.9993
Airplane	7.9993	7.9994	7.9993	7.9993
RSimage1	7.9998	7.9997	7.9997	7.9997
RSimage2	7.9997	7.9997	7.9998	7.9997

Table 11. Information entropy test results.

$$H = -\sum_{i=0}^{255} P(i)\log_2 P(i). \tag{28}$$

As shown in Table 11, after encryption, the information entropy of all images exceeds 7.999 and is close to the theoretical value of 8, indicating that the pixel distributions of the encrypted images are uniform and that the algorithm has strong resistance to statistical attacks.

Image	*R*	G	B	Average
Ours	7.9972	7.9974	7.9972	7.9973
Ref[51]	7.9965	7.9963	7.9967	7.9965
Ref[52]	7.9967	7.9931	7.9939	7.9946
Ref[53]	7.9974	7.9969	7.9973	7.9972
Ref[54]	7.9977	7.9969	7.9969	7.9972
Ref[55]	7.9973	7.9967	7.9967	7.9969

Table 12. Comparison of information entropy results for different methods.

	Original image			Encryption image		
	Horizontal	Vertical	Diagonal	Horizontal	Vertical	Diagonal
	R channel					
RSimage1	0.9187	0.9223	0.8431	− 0.0004	− 0.0002	− 0.0015
Baboon	0.9231	0.8660	0.8543	− 0.0015	− 0.0014	0.0020
Pepper	0.9635	0.9663	0.9564	− 0.0011	0.0035	− 0.0020
Airplane	0.9726	0.9568	0.9343	0.0037	0.0020	0.0011
RSimage2	0.9893	0.9817	0.9730	0.0001	− 0.0022	0.0012
	G channel					
RSimage1	0.9221	0.9257	0.8498	− 0.0007	− 0.0004	− 0.0016
Baboon	0.8655	0.7650	0.7348	0.0035	− 0.0001	0.0024
Pepper	0.9811	0.9818	0.9687	− 0.0010	0.0011	− 0.0010
Airplane	0.9578	0.9678	0.9326	0.0062	0.0020	0.0029
RSimage2	0.9893	0.9817	0.9730	0.0001	− 0.0022	0.0012
	B channel					
RSimage1	0.9804	0.9134	0.8238	0.0004	− 0.00001	0.0007
Baboon	0.9073	0.8809	0.8399	0.0007	0.0036	− 0.0022
Pepper	0.9665	0.9664	0.9478	0.00004	0.0020	0.0013
Airplane	0.9640	0.9353	0.9146	0.0036	0.0019	0.0075
RSimage2	0.9901	0.9830	0.9749	0.0002	− 0.0021	0.0012

Table 13. Statistical results of correlation coefficients.

To further demonstrate the superiority of the proposed encryption algorithm, the Baboon (256 × 256) image was encrypted and compared with methods proposed in other studies. The comparison results are presented in Table 12.

As shown in Table 12, the information entropy achieved by the proposed encryption algorithm is superior to other comparison algorithms, demonstrating the effectiveness of the proposed algorithm and confirming that its information entropy meets the requirements of mainstream encryption algorithms.

Pixel correlation analysis

Pixel correlation is one of the key metrics for evaluating the security of image encryption algorithms. For natural images, adjacent pixels in the horizontal, vertical, or diagonal directions typically exhibit strong correlation. A highly secure encryption algorithm should effectively break this correlation, reducing the correlation coefficients between adjacent pixels in the encrypted image to near zero.

By randomly selecting pairs of adjacent pixels in the horizontal, vertical, and diagonal directions in the encrypted image, the correlation coefficients can be calculated using Eq. (29).

$$r = \frac{Cov(x, y)}{\sqrt{D(x)}\sqrt{D(y)}}. \tag{29}$$

In Eq. (23), x and y represent adjacent pixel pairs, Cov(x, y) denotes the covariance, and D(x) represents the variance. The correlation coefficient results for different directions are presented in Table 13.

As can be seen from Table 13, the correlation between adjacent pixels in the encrypted images is close to 0, demonstrating that the proposed encryption algorithm can effectively eliminate pixel correlation, thereby enhancing the security of the encrypted images.

Chi-square test

The chi-square test is a statistical method used to evaluate whether the pixel value distribution of an encrypted image approximates a uniform distribution. Ideally, an encrypted image should exhibit a uniform distribution of

Image	*R* channel	G channel	B channel
RSimage1	251.84	276.30	210.92
Baboon	276.10	285.17	234.86
Pepper	256.32	247.87	227.73
Airplane	252.53	235.94	264.77
RSimage2	226.05	233.67	270.26

Table 14. Chi-square test results.

image	*R* channel	G channel	B channel	overall
RSimage1	99.6081	99.6098	99.6131	99.6104
Baboon	99.6391	99.6208	99.6120	99.6240
Pepper	99.6078	99.6021	99.6136	99.6078
Airplane	99.6151	99.6162	99.6120	99.6145
RSimage2	99.6034	99.6212	99.6200	99.6149

Table 15. NPCR test results.

pixel values in the range 0–255 to resist statistical attacks. The chi-square test quantifies the degree of deviation between the actual pixel distribution and the theoretical uniform distribution.

For single-channel images with pixel values ranging from 0 to 255, the chi-square value χ^2 is calculated using Eq. (30). In Eq. (30), O_i denotes the number of pixels whose gray level equals *i*, and E_i represents the expected frequency of each gray level under an ideal uniform distribution.

$$\chi^2 = \sum_{i=0}^{256} \frac{(O_i - E_i)^2}{E_i}. \tag{30}$$

At a significance level of α = 0.05, when $\chi^2 \leqslant 293.2478$, the pixel distribution of the image is considered uniform[56]. The chi-square test results for the proposed encryption algorithm are shown in Table 14.

As shown in Table 14, the encrypted images processed by the proposed algorithm exhibit uniform pixel distributions.

NPCR and UACI analysis

To evaluate the proposed encryption algorithm's resistance to differential attacks, NPCR (Number of Pixels Change Rate) and UACI (Unified Average Changing Intensity) metrics are used to test the sensitivity of the encrypted images. For an image of size M × N, NPCR and UACI are calculated using Eqs. (31)–(33), respectively.

$$NPCR = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} T(m,n). \tag{31}$$

$$T(m,n) = \begin{cases} 1, \, if \; \mathrm{V}_1(m,n) \neq V_2(m,n) \\ 0, \, if \; \mathrm{V}_1(m,n) = V_2(m,n). \end{cases} \tag{32}$$

$$UACI = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{[V_1(m,n) - V_2(m,n)]}{255}. \tag{33}$$

In Eqs. (31)–(33), $V_1(m,n)$ and $V_2(m,n)$ represent the pixel values of the images before and after encryption, respectively. NPCR and UACI tests were conducted on five images, and the results are shown in Tables 15 and 16, respectively.

Using the Baboon (256 × 256) image, a comparison between the proposed method and other methods was performed, with the results presented in Table 17.

For 8-bit grayscale or color images, the theoretical NPCR is 99.6094% and the theoretical UACI is 33.4635%. It is generally accepted that NPCR should exceed 99.5% and UACI should be close to the theoretical value (around 33%) for an algorithm to demonstrate strong resistance to differential attacks.

The experimental results in Tables 15, 16 and 17 show that the NPCR of the images encrypted by the proposed algorithm all exceed 99.6% and the UACI values are close to the theoretical value of 33.4635%, outperforming or meeting the security standards of mainstream encryption algorithms. This indicates that the proposed method exhibits excellent sensitivity and security.

image	*R* channel	G channel	B channel	overall
RSimage1	33.5055	33.5069	33.4958	33.5027
Baboon	33.4325	33.4321	33.4326	33.4324
Pepper	33.4554	33.5057	33.4364	33.4658
Airplane	33.5195	33.4607	33.5014	33.4938
RSimage2	33.4903	33.4881	33.4852	33.4879

Table 16. UACI test results.

	NPCR			UACI		
	R channel	G channel	B channel	*R* channel	G channel	B channel
Ours	99.6201	99.6307	99.6353	33.4094	33.4108	33.4760
Ref[52]	99.5809	99.5992	99.5975	33.4076	33.1655	33.2769
Ref[55]	99.5880	99.5880	99.5880	33.4273	33.4635	33.7951
Ref[57]	99.6002	99.6017	99.6002	29.7053	28.0979	30.8363

Table 17. Comparison of NPCR and UACI results for different algorithms.

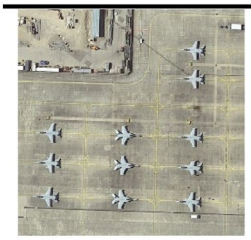
Original Image	Noise0.005	Noise 0.05	Noise 0.1
			

Table 18. Decryption results under salt-and-pepper noise attack.

Robustness analysis

To verify the robustness of the proposed encryption algorithm, this study performed salt-and-pepper noise attack and cropping attack tests on encrypted images, and analyzed the recoverability and anti-interference capability of the decrypted images.

First, in the salt-and-pepper noise attack test, noise with intensities of 0.001, 0.005, and 0.1 was added to the encrypted images. The attacked images were then decrypted, and the subjective visual quality of the decrypted images was examined. The experimental results are shown in Table 18. The results indicate that even under relatively high noise density, the main content of the decrypted images remains recognizable, and the noise points are uniformly distributed, demonstrating good resistance to noise interference.

Next, in the cropping attack test, encrypted images were partially cropped at areas of 6.25%, 25%, and 50%, and the cropped encrypted images were then decrypted. The results are presented in Table 19. As shown in Table 19, despite the information loss, the visible structures of the decrypted images are still well preserved, the overall recognizability remains high, and the noise points are uniformly distributed, proving that the algorithm has strong resistance to cropping attacks.

Cryptographic attack testing

This study conducts chosen-plaintext attack (CPA) analysis and known-plaintext attack (KPA) analysis on the proposed encryption method. The experimental results are shown in Tables 20 and 21. The KPA experiment yields an MSE of 8096.31 and a PSNR of 9.05 dB between P2 and P2_hat.

As shown in Table 20, in the chosen-plaintext attack (CPA) experiments, the ciphertexts generated from typical plaintexts—including all-zero, all-255, single-pixel, and checkerboard patterns—exhibit uniformly distributed noise without any visible structure or plaintext-related features. The corresponding difference images also appear as random noise. Although the ciphertext difference produced by the single-pixel perturbation is relatively weak—indicating that the diffusion strength under extremely small perturbations is limited—the proposed dual-level encryption framework effectively eliminates this minor structural deficiency in the diffusion stage. Overall, the CPA results demonstrate strong confusion and diffusion properties.

According to Table 21, in the known-plaintext attack (KPA) experiment, the attacker constructs an equivalent key stream using one plaintext–ciphertext pair. However, the visualized equivalent key stream still appears as

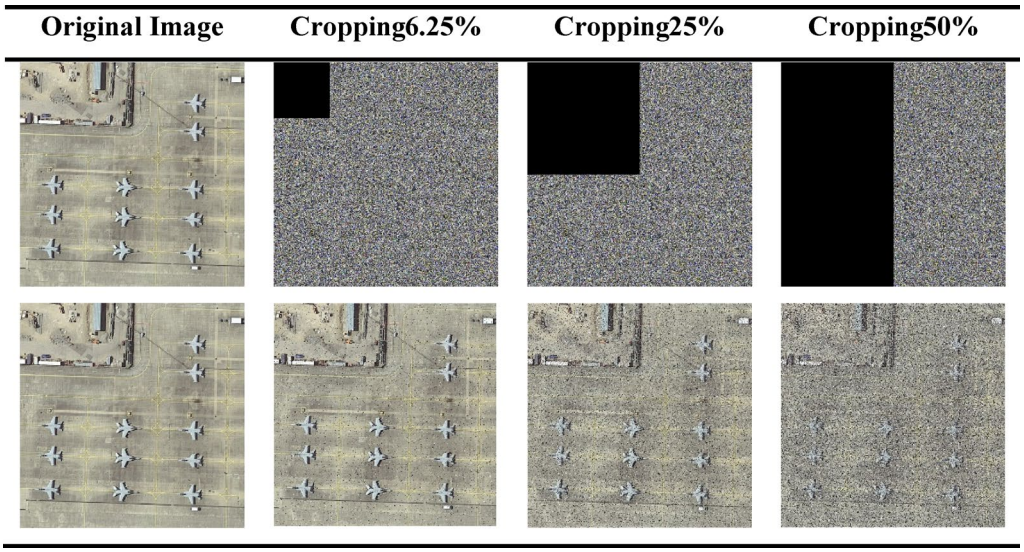


Table 19. Decryption results under cropping attack.

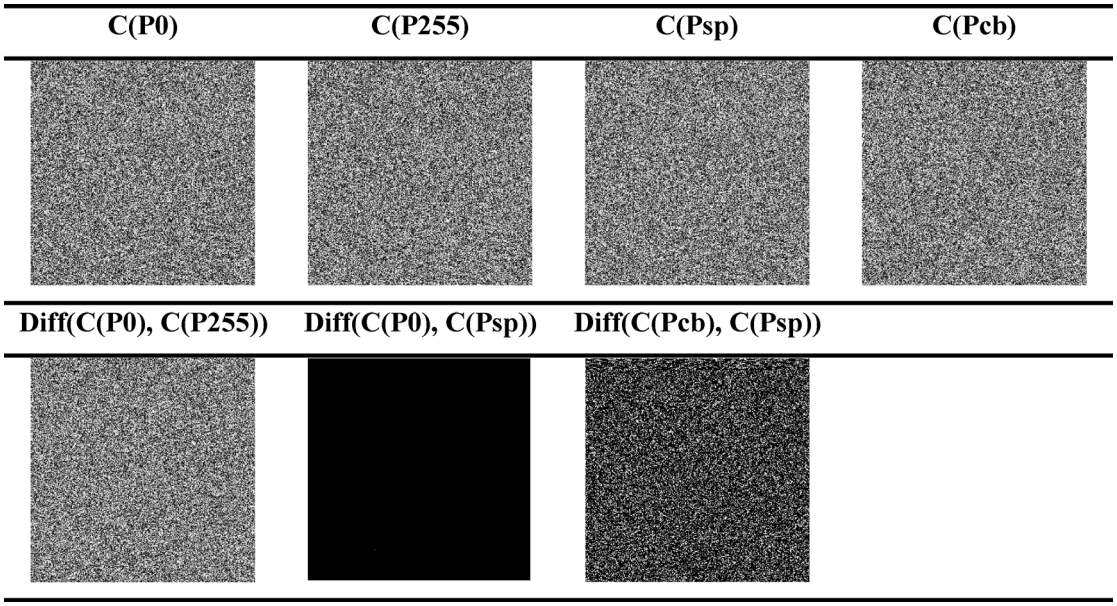


Table 20. Results of the chosen-plaintext attack (CPA).

random noise. Furthermore, when using this key stream to recover another ciphertext, the output is a completely meaningless noise image with no resemblance to the true plaintext.

Combining the results of both attack scenarios, it can be concluded that the proposed algorithm does not leak any structural information in the ciphertext and does not allow valid key extraction from plaintext–ciphertext pairs, thus exhibiting strong security against both CPA and KPA.

A framework for sensitive target detection and encryption

In this work, the improved YOLOv7-tiny is combined with a modified Josephus-ring-based chaotic encryption algorithm to construct a security framework for military aircraft detection and encryption in remote sensing images. The operational flow of the framework is illustrated in Fig. 13.

Although the proposed method further improves the detection accuracy of YOLOv7-tiny for military aircraft in remote sensing scenes, misdetections and missed detections may still occur in practical applications. In the case of a missed detection, the corresponding aircraft target would not be encrypted; conversely, a false detection would lead to unnecessary consumption of encryption resources. However, in real deployment scenarios where the detected targets are of high importance, the proposed detection module is intended to function as an auxiliary tool in conjunction with manual verification. Any missed or incorrect detections can be manually

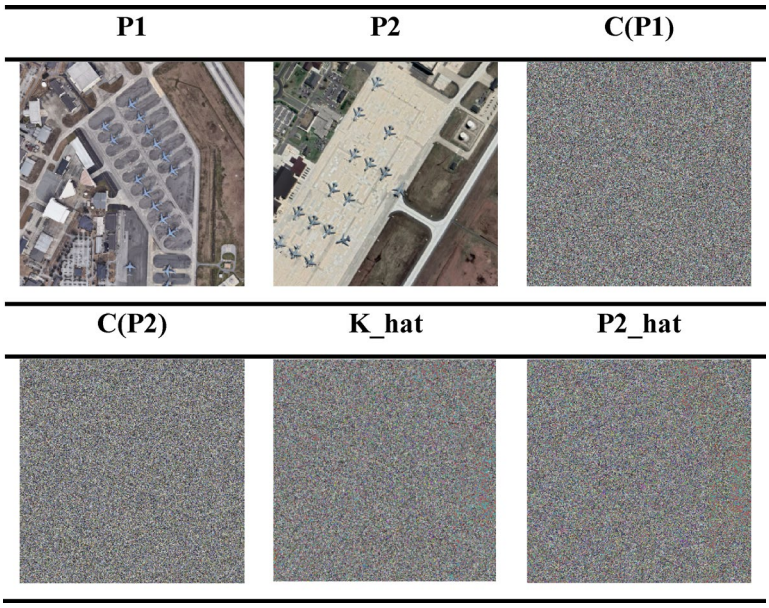


Table 21. Results of the known-plaintext attack (KPA).

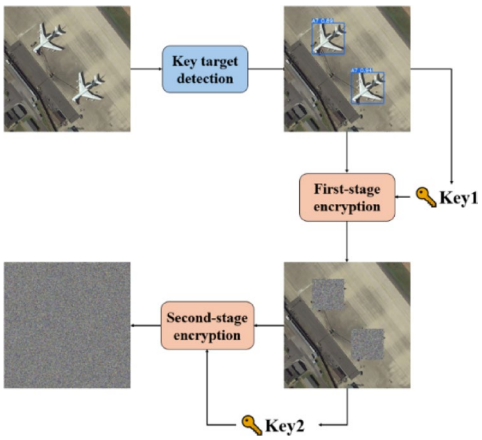


Fig. 13. Overall operational flow of the proposed framework.

corrected by authorized personnel at higher security levels. After this verification stage, secondary encryption is applied, and the processed data are subsequently distributed to lower-level departments. Therefore, potential misdetection or missed-detection issues can be effectively mitigated through human intervention in actual operational workflows.

At the application level, only organizations with high-level security clearance are granted simultaneous access to Key1 and Key2, enabling full-resolution viewing of the image, including sensitive target details. In contrast, lower-level security units are permitted to access only Key2, which ensures that they can analyze the encrypted images while preventing the disclosure of sensitive aircraft information embedded within them.

If a lower-level unit requires access to sensitive target details for operational purposes, it may submit a request to the higher-level authority that originally issued the data. Upon approval, Key1 can be provided in a controlled manner, thereby maintaining strict information security throughout the entire workflow.

Conclusion

At present, research on the integration of sensitive-target detection and localized encryption in remote sensing imagery remains limited, and there is still room for exploration regarding hierarchical data access within a unified security framework. To address these gaps, this paper proposes a security framework for military aircraft detection combined with dual-level encryption. By integrating a YOLO-based object detection algorithm with a chaos-driven encryption scheme, the proposed framework achieves high-precision detection and classification of military aircraft while ensuring sensitive information protection and enabling hierarchical data access control.

To enhance detection accuracy under complex remote sensing backgrounds, the YOLOv7-tiny model is adopted and improved through the introduction of a novel Multi-branch Enhanced Feature Aggregation Block (MEFABlock). MEFABlock employs a three-branch convolutional architecture with different kernel sizes to capture multi-scale receptive fields. After convolutional operations, an Efficient Channel Attention (ECA) mechanism is incorporated to strengthen the perception of key features, and a Convolutional Block Attention Module (CBAM) is embedded at the output to optimize both channel-level and spatial-level feature representations. These enhancements collectively improve the model's sensitivity to aircraft targets in challenging remote sensing environments. In addition, coordinate convolution is integrated into the detection head to further enhance spatial modeling capability. Through the synergy of these modules, the improved detector achieves notable performance gains. Compared with the baseline, Precision (P) increases by 2.4%, Recall (R) by 2.7%, mAP@0.5 by 2.7%, and mAP@0.5:0.95 by 2.2%.

Furthermore, a chaotic encryption algorithm consisting of pixel diffusion, an improved Josephus-ring-based permutation, and a second diffusion stage is proposed. The encryption procedure is designed hierarchically: detected aircraft regions are first encrypted locally, followed by global encryption of the entire image to reinforce overall security. Different hash-based keys are used in the two encryption stages to achieve hierarchical access control, allowing users of different authorization levels to decrypt content according to their permissions. A novel two-dimensional multi-trigonometric chaotic map (2D-MTCM) is developed as the chaotic system, along with a hash-based method for computing its initial parameters. Chaos characteristics of 2D-MTCM are validated through Lyapunov exponent analysis, bifurcation diagrams, and the 0–1 test, demonstrating its favorable dynamical properties. Experimental results show that the encrypted remote sensing images achieve an information entropy of 7.9997. Additional experiments—including pixel correlation and robustness evaluations—further confirm the strong encryption performance and attack resistance of the proposed algorithm.

The proposed method employs an improved lightweight YOLOv7-tiny model, which enables effective detection of military aircraft targets while meeting real-time processing requirements. However, considering the hardware constraints commonly encountered in practical engineering deployments, there is still room to further reduce the computational cost of the model. In addition, chaos evaluations indicate that the chaotic behavior of the 2D-MTCM system can be further enhanced, and the CPA results reveal minor structural weaknesses in the diffusion stage of the encryption algorithm. Future work will therefore focus on optimizing the encryption process to eliminate diffusion deficiencies and enhance the chaotic dynamics of the system. Moreover, additional research will be conducted to further lightweight the detection model to better suit resource-constrained deployment scenarios.

Data availability

The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Received: 11 October 2025; Accepted: 15 December 2025
Published online: 24 December 2025

References

1. Kong, L., Wang, J. & Zhao, P. YOLO-G: A lightweight network model for improving the performance of military targets detection. *IEEE Access.* **10**, 55546–55564 (2022).
2. Sun, Y. et al. YOLO-E: A lightweight object detection algorithm for military targets. *Signal. Image Video Process.* **19**, 241 (2025).
3. Ren, S., He, K., Girshick, R., Sun, J. & Faster, R-C-N-N. Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149 (2017).
4. Dai, J., Li, Y., He, K. & Sun, J. R. F. C. N. Object detection via region-based fully convolutional networks. In *Adv Neural Inf. Process. Syst* 379–387 (2016).
5. Nelson, J. & Solawetz, J. YOLOv5 is here: State-of-the-art object detection at 140 FPS. Preprint at https://ultralytics.com (2020).
6. Li, C. et al. YOLOv6: A single stage object detection framework for industrial applications. Preprint at (2022). https://arxiv.org/abs/2209.02976.
7. Wang, C. Y., Bochkovskiy, A. & Liao, H. Y. M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Preprint at (2022). https://arxiv.org/abs/2207.02696.
8. Jocher, G., Chaurasia, A. & Qiu J. Ultralytics YOLO. *Zenodo*. https://doi.org/10.5281/zenodo.13942966 (2023).
9. Wang, C. Y., Yeh, I. H. & Liao, H. Y. M. YOLOv9: Learning what you want to learn using programmable gradient information. Preprint at (2024). https://arxiv.org/abs/2402.13616.
10. Wang, A. et al. YOLOv10: Real-time end-to-end object detection. Preprint at (2024). https://arxiv.org/abs/2405.14458.
11. Jocher, G. & Qiu, J. Ultralytics YOLO11. (2024). https://github.com/ultralytics/yolov11.
12. Tian, Y., Ye, Q. X. & Doermann, D. YOLOv12: Attention-centric real-time object detectors. Preprint at (2025). https://arxiv.org/abs/2502.12524.
13. Anonymous. YOLOv13: Real-time object detection with hypergraph-enhanced adaptive visual perception. Preprint at (2025). https://arxiv.org/abs/2506.17733.
14. Wang, X. et al. AG-YOLO: Attention-guided YOLO for efficient remote sensing oriented object detection. *Remote Sens.* **17**, 1027 (2025).
15. Wu, T., Dong, Y. & YOLO-SE Improved YOLOv8 for remote sensing object detection and recognition. *Appl. Sci.* **13**, 12977 (2023).
16. Hong, B., Li, W., Zhu, W., Wang, X. & Zhang, K. Research on target detection based on improved yolov3-spp algorithm. *J. Ordnance Equip. Eng.* **44**, 1–8 (2023).
17. Liu, K. et al. MOKP-YOLO: A unified high-performance model for military object and key part detection in UAV images. *J. Supercomput.* **81**, 1–29 (2025).
18. Xi, L. H. et al. A multiscale information fusion network based on pixel-shuffle integrated with YOLO for aerial remote sensing object detection. *IEEE Geosci. Remote Sens. Lett.* **21**, 1–5 (2024).
19. Liu, K. et al. Military aircraft recognition method based on attention mechanism in remote sensing images. *IET Image Process.* **19**, e70069 (2025).
20. Xi, Y. et al. Structure-adaptive oriented object detection network for remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **62**, 13 (2024).

21. Shi, T. et al. Progressive class-aware instance enhancement for aircraft detection in remote sensing imagery. *Pattern Recognit.* **164**, 111503 (2025).
22. Li, J., Yang, N., Li, M. & Li, S. Medical image encryption based on a novel four-dimensional chaotic system. *Sci. Technol. Eng.* **25**, 1–8 (2025).
23. Li, H. et al. Exploiting dynamic vector-level operations and a 2D-enhanced logistic modular map for efficient chaotic image encryption. *Entropy* **25**, 1147 (2023).
24. Yu, F. et al. Dynamic analysis and application of 6D multistable memristive chaotic system with wide range of hyperchaotic States. *Axioms* **14**, 638 (2025).
25. Yu, F. et al. Dynamics analysis, synchronization and FPGA implementation of multiscroll Hopfield neural networks with non-polynomial memristor. *Chaos Solitons Fract.* **179**, 114440 (2024).
26. Yu, F. et al. Bursting firings in memristive Hopfield neural network with image encryption and hardware implementation. *IEEE Trans. Comput. -Aided Des. Integr. Circuits Syst.* **44**, 4564–4576 (2025).
27. Ye, C. et al. Social image security with encryption and watermarking in hybrid domains. *Entropy* **27**, 276 (2025).
28. Wang, G., Ye, X. & Zhao, B. A novel remote sensing image encryption scheme based on block period Arnold scrambling. *Nonlinear Dyn.* **112**, 17477–17507 (2024).
29. Teng, L. et al. Chaotic image encryption based on partial face recognition and DNA diffusion. *Appl. Intell.* **54**, 10360–10373 (2024).
30. Kumar, S. & Sharma, D. Image scrambling encryption using chaotic map and genetic algorithm: A hybrid approach for enhanced security. *Nonlinear Dyn.* **112**, 12537–12564 (2024).
31. Odeh, A. et al. Lightweight secure image encryption: A tent map chaos theory approach. *Multimed Tools Appl.* –, 1–20 (2025).
32. Maity, A. & Dhara, B. C. Image encryption utilizing 5D hyperchaos, wavelet lifting scheme, and Burrows–Wheeler transform. *Arab. J. Sci. Eng.* –, 1–17 (2025).
33. Zhang, S. & Liu, L. A novel image encryption algorithm based on SPWLCM and DNA coding. *Math. Comput. Simul.* **190**, 723–744 (2021).
34. Feng, W. et al. Cryptanalysis and improvement of the image encryption scheme based on feistel network and dynamic DNA encoding. *IEEE Access.* **9**, 145459–145470 (2021).
35. Wen, H. & Yu, S. Cryptanalysis of an image encryption cryptosystem based on binary bit planes extraction and multiple chaotic maps. *Eur. Phys. J. Plus.* **134**, 337 (2019).
36. Feng, W. et al. Security analysis of an image encryption algorithm based on variable step length Josephus traversing and DNA dynamic encoding. *J. Electron. Inf. Technol.* **44**, 3635–3642 (2022).
37. Wen, H. et al. Security analysis of a color image encryption algorithm using a fractional-order chaos. *Entropy* **23**, 258 (2021).
38. Wang, Q. et al. ECA-Net: Efficient channel attention for deep convolutional neural networks. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* 11534–11542 (2020).
39. Woo, S., Park, J., Lee, J. Y. & Kweon, I. S. CBAM: Convolutional block attention module. In *Proc. Eur. Conf. Comput. Vis. (ECCV)* 3–19 (2018).
40. Liu, R et al. An intriguing failing of convolutional neural networks and the CoordConv solution. Preprint at. (2018). https://doi.org/10.48550/arXiv.1807.03247.
41. Sun, S. et al. A color image encryption scheme utilizing a logistic-sine chaotic map and cellular automata. *Sci. Rep.* **15**, 21603. https://doi.org/10.1038/s41598-025-04968-4 (2025).
42. Wang, X. et al. A color image encryption and hiding algorithm based on hyperchaotic system and discrete cosine transform. *Nonlinear Dyn.* **111**, 14513–14536 (2023).
43. Huang, X., Tang, J. & Zhang, Z. Efficient and secure image encryption algorithm using 2D LIM map and Latin square matrix. *Nonlinear Dyn.* **112**, 22463–22483 (2024).
44. Hua, Z., Zhang, Y. & Zhou, Y. Two-dimensional modular chaotification system for improving chaos complexity. *IEEE Trans. Signal. Process.* **68**, 1937–1949. https://doi.org/10.1109/TSP.2020.2979596 (2020).
45. Hua, Z. et al. Color image encryption using orthogonal Latin squares and a new 2D chaotic system. *Nonlinear Dyn.* **104**, 4505–4522. https://doi.org/10.1007/s11071-021-06472-6 (2021).
46. Nan, S. X. et al. Remote sensing image compression and encryption based on block compressive sensing and 2D-LCCCM. *Nonlinear Dyn.* **108**, 2705–2729. https://doi.org/10.1007/s11071-022-07335-4 (2022).
47. Yang, N. et al. Medical image encryption based on Josephus traversing and hyperchaotic Lorenz system. *J. Shanghai Jiaotong Univ. (Sci).* **29**, 91–108 (2024).
48. Yu, W. et al. Mar20: A dataset for military aircraft target recognition in remote sensing images. *J. Remote Sens.* **27**, 2688–2696 (2023).
49. Zhao, Y. et al. June. DETRs beat YOLOs on real-time object detection. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* 16965–16974 (Seattle, WA, USA, 2024).
50. Ren, Q., Teng, L., Wang, X. Y. & Jiang, D. H. A visually secure image encryption scheme based on compressed sensing and Chebyshev-dynamics coupled map lattices in cloud environment. *Eur. Phys. J. Plus.* **138**, 1–20 (2023).
51. Alexan, W. et al. Color image encryption through Chao and KAA map. *IEEE Access* **11**, 11541–11554 (2023).
52. Tanveer, M. et al. Multi-images encryption scheme based on 3D chaotic map and substitution box. *IEEE Access* **9**, 73924–73937 (2021).
53. ul Haq, T. & Shah, T. 4D mixed chaotic system and its application to RGB image encryption using substitution–diffusion. *J. Inf. Secur. Appl.* **61**, 102931 (2021).
54. Abd El-Latif, A. A., Abd-El-Atty, B. & Venegas-Andraca, S. E. Controlled alternate quantum walk-based pseudo-random number generator and its application to quantum color image encryption. *Phys. A.* **547**, 123869 (2020).
55. Alexan, W., ElBeltagy, M. & Aboshousha, A. RGB image encryption through cellular automata, S-box and the Lorenz system. *Symmetry* **14**, 443 (2022).
56. Geng, S. et al. Image encryption scheme based on Thorp shuffle and Pseudo dequeue. *Sci. Rep.* **15**, 11141 (2025).
57. Youssef, M. et al. Enhancing satellite image security through multiple image encryption via hyperchaos, SVD, RC5, and dynamic S-box generation. *IEEE Access* **12**, 123921–123945 (2024).

Author contributions

H. W., X. L. and S. Z. contributed to the conceptualization and supervision of the study. They provided theoretical guidance, critical feedback, and were responsible for reviewing and revising the manuscript at all stages. Q. Z. and J. L. were responsible for the development and implementation of the proposed algorithm, including code writing, data collection, investigation, methodology, and drafting the original version of the manuscript.

Funding

This research was funded by the Jilin Provincial Science and Technology Development Program, grant number No. SKL202402023.

Declarations

Competing interests
The authors declare no competing interests.

Consent for publication
All authors have read and approved the final version of the manuscript and consent to its submission and publication.

Additional information
Correspondence and requests for materials should be addressed to H.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.