Computer Standards & Interfaces 97 (2026) 104115

Contents lists available at ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi

Robust zero-watermarking method for multi-medical images based on Chebyshev–Fourier moments and Contourlet-FFT

Xinhui Lu [a], Guangyun Yang [a], Yu Lu [a], Xiangguang Xiong [a,b,*]

[a] School of Big Data and Computer Science, Guizhou Normal University, Guiyang 550025, China
[b] Guizhou Provincial Specialized Key Laboratory of Information Security Technology in Higher Education Institutions, Guiyang, 550025, China

ARTICLE INFO

Keywords:
Zero-watermarking
Lorenz chaotic system
Chebyshev–Fourier moments
Contourlet transform
Fast Fourier transform

ABSTRACT

Classical robust watermarking methods embed secret data into a cover image designed to protect its copyright. However, they suffer from the problem of balancing imperceptibility and robustness. To address this issue, the impact of conventional attacks on the stability of feature vectors extracted from the cover image is examined. Accordingly, we proposed a zero-watermarking method with high attack resistance for multi-medical images by employing Contourlet transform (CT), Chebyshev–Fourier moments (CHFMs), and fast Fourier transform (FFT). First, each medical image is normalized separately, and the normalized images are fused using a dual-tree complex wavelet transform-based method. Second, the effective region is extracted and subjected to the CT. The CHFMs of the low-frequency sub-bands are calculated, and the FFT is performed on the generated amplitude sequence to construct a feature matrix. A feature image is generated by combining the magnitude of each feature value with the overall mean. Finally, the copyrighted image is encrypted using the Lorenz chaotic system and Fibonacci Q-matrix, after which an exclusive-OR operation is applied between the generated feature image and the encrypted copyrighted image to produce a zero-watermarking signal. The results show that the proposed method exhibits excellent resistance to attack with a normalized correlation coefficient of up to 0.994 between the extracted image and the original copyrighted one. Furthermore, the average anti-attack performance of our proposed method is approximately 2% higher compared to similar existing methods, indicating that our proposed method is highly resistant to conventional, geometric, and combinatorial attacks.

1. Introduction

Steganography is a widely used technique for covertly embedding secret data within multimedia covers, aiming to ensure undetectability and robustness. By effectively concealing data presence, it enhances security and privacy, with broad applications across various fields [1–4]. Unlike steganography, which has the primary purpose of concealing the existence of data, robust digital watermarking techniques [5–8] aim to confirm copyright ownership by embedding specific secret data in the protected object. However, because of the strategy used to embed the secret data into the cover, increasing the strength of the embedding degrades the quality of the cover, thus damaging cover integrity. To address the limitations of traditional watermarking methods, Wen et al. [9] introduced zero-watermarking. Unlike conventional approaches, this technique preserves the original image by generating authentication data from stable image features rather than altering pixel values, ensuring both integrity and copyright protection. As a result, the zero-watermarking technique can effectively balance the contradiction between reducing the original cover image's quality and ensuring the robustness and imperceptibility of traditional embedded watermarking techniques, making it valuable for essential applications in many fields, such as multimedia data management.

With the different domains used in constructing feature images, there are three categories of zero-watermarking techniques. The first type comprises spatial-domain-based zero-watermarking methods [10–15]. Yang et al. [10] suggested a zero-watermarking method that uses the center pixels of different channels of the cover image as the center of the circle, whereby the pixels covered by rings with different radii and widths constitute the feature image. After that, the final zero-watermarking signal is generated by executing an exclusive-OR operation on the encrypted copyrighted and feature images. Chang et al. [11] proposed a method using secret sharing exhibiting strong robustness and security. Chang et al. [12] used a Sobel operator to extract the texture and edge features of a cover image to construct a robust zero-watermarking signal. Zou et al. [13] proposed a similarity retrieval method with good resistance to attack. These methods process

* Corresponding author at: School of Big Data and Computer Science, Guizhou Normal University, Guiyang 550025, China.
E-mail address: xxg0851@163.com (X. Xiong).

https://doi.org/10.1016/j.csi.2025.104115
Received 3 April 2025; Received in revised form 12 November 2025; Accepted 8 December 2025
Available online 8 December 2025

the cover image directly in the pixel domain, offering implementation advantages of simplicity and intuitiveness.

The second type of method is frequency-domain-based [16–22]. Yang et al. [16] proposed a method that was based on the non-subsampled Shearlet transform and Schur decomposition, which achieved better anti-attack performance. Huang et al. [17] extracted low-frequency sub-bands (LSs) using a dual-tree complex wavelet transform (DTCWT), partitioned the LSs, and used Hessenberg decomposition to yield a robust signal. Lu et al. [18] proposed fusing the cover images with the gray-weighted averaging fusion method, generating a robust zero-watermarking signal using the fast finite Shearlet transform and Schur decomposition. Wu et al. [20] presented a robust scheme for constructing a zero-watermarking signal to encrypt medical images using the Contourlet transform (CT). These methods generate robust zero-watermarking signals by transforming the cover image from the spatial to the frequency domain, leveraging frequency-domain properties that offer enhanced resistance against non-geometric attacks.

Although the above methods are effective against conventional image processing attacks, they are not against large-scale geometric attacks such as rotation, scaling, and cropping, because the features extracted by these methods are not geometrically invariant. Therefore, to enhance the resilience of zero-watermarking methods against geometric attacks, some scholars have proposed the use of continuous orthogonal moments that possess stability and geometric invariance [23–27], to optimize the construction and verification of zero-watermarking signals. This falls into the third category of zero-watermarking methods. Bessel–Fourier moments (BFMs) [23] are among the most representative continuous orthogonal moments. Their radial polynomials are considered to be feature functions with good orthogonality and are widely used in the field of pattern recognition. Gao et al. [24] proposed a robust method using BFMs. This method first normalizes the cover image to get translation and scaling invariance. Then it computes the BFMs of the normalized image to construct a zero-watermarking signal using moment–rotation invariance. Subsequently, neural networks were introduced into watermarking techniques to comprehensively improve their adaptability and robustness in the face of complex and changing image processing and geometric attacks [28–36]. Gong et al. [28] proposed a robust medical image zero-watermarking method based on a residual DenseNet. He et al. [29] proposed a robust image method based on shrinkage and a redundant feature elimination network. Such methods provide a higher level of understanding and protection of image content using the superb feature extraction capability of Neural networks. However, Neural network-based zero-watermarking methods face multiple challenges, including substantial training data requirements, high computational complexity, limited interpretability, and susceptibility to adversarial attacks.

All of the above methods satisfy the basic requirements of digital watermarking technology. However, most of these methods lack a strong anti-attack ability to resist diverse attacks, with poor performance against geometric and combinatorial attacks. Additionally, the costs of centralized protection and the occupation of storage space for multiple images are relatively high. To address these issues, a zero-watermarking method that combines CT, Chebyshev–Fourier moments (CHFMs), and fast Fourier transform (FFT) is proposed. This approach leverages the directional selectivity and sparsity of CT, the orthogonality and rotational invariance properties of CHFMs, and the computationally efficient and numerically stable properties of FFT. Compared with zero-watermarking methods that only use frequency domain or orthogonal moments, this method enhances robustness against geometric and combinatorial attacks by combining CT and CHFMs, which fully utilize the multi-scale features of CT and the geometric invariance of CHFMs. Additionally, the method adopts DTCWT-based fusion for efficient multi-image protection and storage reduction. The main contributions are as follows:

(1) The effects of conventional attacks on the stability of feature vectors extracted from cover images were analyzed. The results indicate that the extracted feature vectors are highly resistant to attacks.

(2) The protection cost of multiple medical images was reduced by a fusion operation using a dual-tree complex wavelet transform-based method.

(3) The CT and CHFMs were employed to construct a zero-watermarking signal to address the problem that existing methods are only resistant to limited attacks.

(4) The Lorenz chaotic system and Fibonacci Q-matrix were utilized to encrypt the copyrighted image to heighten the proposed method's security.

The remainder of the paper is organized as follows: Section 2 presents the basic theory, including the Lorenz chaotic system and Fibonacci Q-matrix, image normalization technique, image fusion method using DTCWT, CHFMs, CT, and FFT. Section 3 analyzes the effect of conventional attacks on the stability of feature vectors extracted from cover images. Section 4 describes the key steps of the copyrighted image encryption, zero-watermarking signal construction, and detection. Section 5 presents the attack resistance of our method and evaluates its superiority by comparing it with similar ones. The final section concludes this paper.

2. Basic theory

2.1. Lorenz chaotic system

The Lorenz chaotic system is a nonlinear dynamic system discovered by the American meteorologist Edward Norton Lorenz in 1963 during his research on weather changes. The system models atmospheric convective motion using three-dimensional ordinary differential equations, generating high-quality chaotic sequences free from short-cycle effects. Its unpredictability and randomness make it particularly suitable for image encryption applications. The Lorenz chaotic system [37] is represented as follows:

$$\begin{cases} \dot{x} = a(y - x) \\ \dot{y} = x(c - z) - y \\ \dot{z} = xy - bz \end{cases} \tag{1}$$

where a, b, and c denote the three constants of the Lorenz chaotic system, and x, y, and z represent its three state variables. The Lorenz chaotic system produces a butterfly-shaped chaotic attractor as displayed in Fig. 1(a) when $a = 10$, $b = \frac{8}{3}$, $c = 28$, and $(x_0, y_0, z_0) = (0.1, 0.1, 0.1)$. In Fig. 1(a), the system is bounded, stochastic, and non-periodic. Fig. 1(b) and (c) illustrate the bifurcation and Lyapunov exponential plots under variations in parameter c.

2.2. Fibonacci Q-matrix

To enhance encryption security and reliability, the researchers utilize properties of the Fibonacci sequence in their method design, significantly improving protection capabilities for enhanced privacy preservation and information security. The recurrence formula for the Fibonacci sequence [38] is calculated as follows:

$$F_n = F_{n-1} + F_{n-2}, n > 2 \tag{2}$$

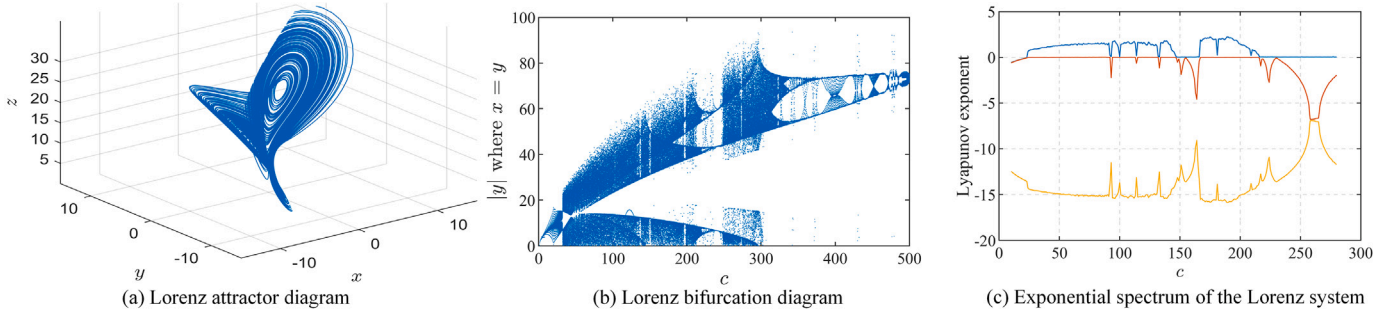
where $F_1 = F_2 = 1$, F_n denotes the nth Fibonacci number.

The Fibonacci Q-matrix is constructed using Fibonacci numbers. It is usually represented as a 2×2 matrix as follows:

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \tag{3}$$

The corresponding inverse matrix Q^{-1} of the Q-matrix is defined as:

$$Q^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix} \tag{4}$$



(a) Lorenz attractor diagram (b) Lorenz bifurcation diagram (c) Exponential spectrum of the Lorenz system

Fig. 1. Chaotic attractor diagram, bifurcation diagram, and Lyapunov exponential spectrum of the Lorenz system.

(a) Heart (b) Chest X-ray (c) Brain (d) Knee (e) Normalized of (a) (f) Normalized of (b) (g) Normalized of (c) (h) Normalized of (d)

Fig. 2. Experiment results of image normalization.

The nth power of the Q-matrix is defined as follows:

$$Q^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix} \tag{5}$$

The determinants of the Q-matrix can be expressed as:

$$\mathrm{Det}(Q^n) = F_{n+1}F_{n-1} - F_n^2 = (-1)^n \tag{6}$$

The corresponding inverse matrix Q^{-n} of the Q^n is given below:

$$Q^{-n} = \begin{bmatrix} F_{n-1} & -F_n \\ -F_n & F_{n+1} \end{bmatrix} \tag{7}$$

2.3. Image normalization

Normalization is a critical step in image processing and computer vision [39]. Generally, the cover image after the normalization operation is transformed into a standard form that can resist attacks from affine transformations, such as translation, rotation, and scaling. The two-dimensional $(p+q)$-order moment of the cover image $f(x,y)$ is defined as:

$$m^{pq} = \sum_x \sum_y x^p y^q f(x,y) \tag{8}$$

where p, $q = 0, 1, 2, 3...$, and the image central moments are defined as

$$u^{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y) \tag{9}$$

where $(\bar{x}, \bar{y})$ is the center of mass of the image with $\bar{x} = \frac{m_{10}}{m_{00}}$ and $\bar{y} = \frac{m_{01}}{m_{00}}$.

The covariance matrix M of the cover image is defined as $\begin{bmatrix} u_{20} & u_{11} \\ u_{11} & u_{02} \end{bmatrix}$. The normalization operation for the image is based on the invariance of the matrix as follows:

$$\begin{bmatrix} x^m \\ y^m \end{bmatrix} = \begin{bmatrix} \cos_\alpha & \sin_\partial \\ -\sin_\partial & \cos_\alpha \end{bmatrix} \begin{bmatrix} \frac{c}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{c}{\sqrt{\lambda_2}} \end{bmatrix} \begin{bmatrix} e_{1x} & e_{1y} \\ -e_{1y} & e_{1x} \end{bmatrix} \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix} \tag{10}$$

where λ_1 and λ_2 are the eigenvalues of M, and the corresponding eigenvectors are $[e_{1x}, e_{1y}]^T$ and $[e_{2x}, e_{2y}]^T$, respectively.

The images before and after the normalization process are shown in Fig. 2 for four standard medical images, where the original images are shown in (a)–(d), and the corresponding normalized versions are shown in (e)–(h).

2.4. Image fusion using DTCWT

The dual-tree complex wavelet transform (DTCWT) is a technique that combines the multi-scale analysis capabilities of the discrete wavelet transform with high computational efficiency. It utilizes a tree structure with low-pass and high-pass filter banks to decompose the real and imaginary parts of the image into multiple scales. At each scale, the DTCWT generates a low-frequency component and six detail components with different orientations (±15°, ±45°, ±75°). Recently, DTCWT has been widely adopted in image fusion [39]. The DTCWT efficiently extracts multi-scale image details, producing fused images with richer content and improved visual quality.

(1) DTCWT of the image. Apply DTCWT to the original images for 1-level decomposition to obtain the low-frequency coefficients (LL_1, LL_2, …, LL_k) and high-frequency coefficients (HL_1, LH_2, …, HH_k,) with the following equations:

$$\left[LL_k, HL_k, LH_k, HH_k\right] = \mathrm{DTCWT}\left(I_k\right) \tag{11}$$

where $k = 1, 2, \ldots, n$.

(2) Fusion of high-frequency coefficients. Calculate the energy of each coefficient and its neighboring region in the high-frequency subbands of all images. The window size is set to $2r+1$, where r is the window radius. Within this window, each coefficient is given a weight of $\frac{1}{(2r+1)^2}$. The local energy $E(x,y)$ of image k at position (x,y) is calculated as:

$$E_k(x,y) = \sum_{m=x-r}^{x+r} \sum_{n=y-r}^{y+r} (|f_k(m,n)|^2) \tag{12}$$

The fused high-frequency coefficients are selected from the image with maximum energy at each position. The relevant formula is shown below.

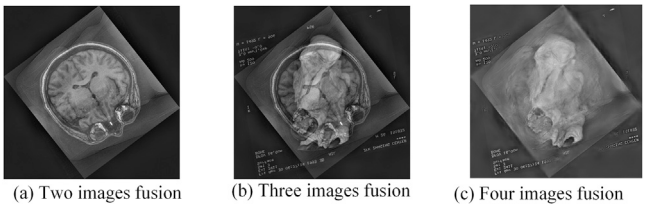
$$HF(x,y) = \arg\max_k E_k(x,y) \tag{13}$$

For positions where multiple images have equal maximum energy, the average of the coefficients of those images is taken.

(3) Fusion of low-frequency coefficients. The maximum coefficients across all images in the LSs are selected.

$$LF(i,j) = \max(LL_1(i,j), LL_2(i,j), \ldots, LL_k(i,j)) \tag{14}$$

(4) Image reconstruction. With the fused high-frequency details and low-frequency coefficients, the reconstructed image is obtained by applying the inverse DTCWT to the fused data.



(a) Two images fusion (b) Three images fusion (c) Four images fusion

Fig. 3. Experimental results of image fusion using DTCWT.

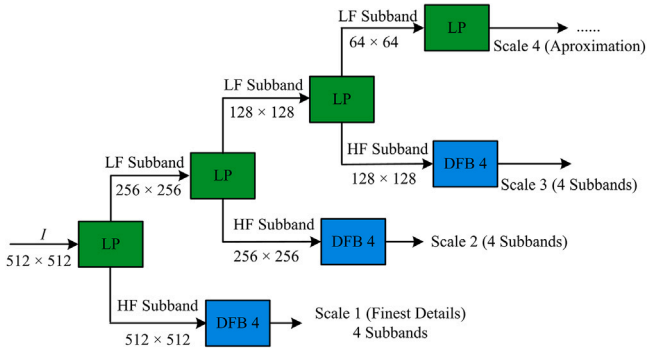


Fig. 4. Schematic of the CT.

The normalized images from Fig. 2 were fused using the aforementioned image fusion technique. Fig. 3 presents the fusion results, where Fig. 3(a) shows the fusion of images Fig. 2(f) and (g); Fig. 3(b) displays the fusion incorporating images Fig. 2(e)–(g); and Fig. 3(c) demonstrates the fusion combining images Fig. 2(e)–(h).

2.5. Contourlet transform

The Contourlet transform (CT) [40] is a dual-filter structure that is effective in obtaining sparse extensions of typical images with smooth contours due to its unique multi-resolution and multidirectional capability. The Laplace Pyramid is utilized to capture point discontinuities in the image, while a bank of directional filters connects these discontinuities into a linear structure. Basic elements such as contour lines are used for image expansion, which facilitates the reconstruction of complex image features. Fig. 4 shows a schematic of the decomposition of a 512 × 512 image using CT.

2.6. Chebyshev-Fourier moments

The Chebyshev–Fourier moments (CHFMs) were proposed by Ping et al. [41] in 2002 and entail the following key steps:

In polar coordinates (r,θ), the Chebyshev–Fourier function $P_{nm}(r,\theta)$ consists of two components: the radial function $R_n(r)$ and the angular function $\exp(jm\theta)$.

$$P_{nm}(r,\theta) = R_n(r)\exp(jm\theta) \tag{15}$$

where

$$R_n(r) = \sqrt{\frac{8}{\pi}}\left(\frac{1-r}{r}\right)^{\frac{1}{4}} \sum_{k=0}^{\frac{n+2}{2}} (-1)^k \frac{(n-k)!}{k!(n-2k)!}[2(2r-1)]^{n-2k} \tag{16}$$

In 2007, Ping et al. [42] showed that CHFMs are deformations of the Jacobi–Fourier moments ($p=2$, $q=3/2$), and thus the radial function $R_n(r)$ of CHFMs can be expressed as

$$R_n(r) = \sqrt{\frac{8}{\pi}}\left(\frac{1-r}{r}\right)^{\frac{1}{4}} \sum_{k=0}^{n} (-1)^k \frac{(n+k+1)!2^{2k}}{(n-k)!(2k+1)!} r^s \tag{17}$$

The functions $P_{nm}(r,\theta)$ are orthogonal within the unit circle, where $(0 \le r \le 1,\ 0 \le \theta \le 2\pi)$.

$$\int_0^{2\pi}\int_0^1 P_{nm}(r,\theta)P_{kl}(r,\theta)rdrd\theta = \delta_{nmkl} \tag{18}$$

where δ_{nmkl} is the Kronecker delta, the image function $f(r,\theta)$ can be decomposed orthogonally in the polar coordinate system by the functional system $P_{nm}(r,\theta)$. Reconstruction using the CHFMs is thus made possible, and the image reconstruction function $f(r,\theta)$ can subsequently be written as:

$$f(r,\theta) = \sum_{n=0}^{\infty}\sum_{m=-\infty}^{+\infty} \phi_{nm} R_n(r)\exp(jm\theta) \tag{19}$$

where ϕ_{nm} is the CHFM for image $f(r,\theta)$.

$$\phi_{nm} = \int_0^{2\pi}\int_0^1 f(r,\theta)R_n(r)\exp(-jm\theta)rdrd\theta \tag{20}$$

2.7. Fast Fourier transform

The FFT is a fast algorithm based on the discrete Fourier transform (DFT) that leverages the inherent properties of the DFT, including symmetry, periodicity, and the relationship between odd and even terms. It works by using its intrinsic periodicity and symmetry to decompose a long sequence of DFTs into the sum of many short sequences of DFTs [43]. The FFT can be represented mathematically as follows:

$$x_k = \sum_{n=0}^{N-1} x_n \bullet e^{-i2\pi k\frac{n}{N}} \tag{21}$$

where $k = 0, 1, 2, \ldots..., N\text{-}1$.

Computing the DFT of a discrete signal using Eq. (21) requires $N \times N$ steps, whereas the FFT computes the DFT of a discrete signal by dividing the DFT equation into two independent components, as shown in Eq. (22).

$$x_k = \sum_{m=0}^{(\frac{N}{2})-1} x_{2m} \bullet e^{-i2\pi k\frac{m}{(N/2)}} + e^{-i2\pi k\frac{1}{N}} \sum_{m=0}^{(\frac{N}{2})-1} x_{2m+1} \bullet e^{-i2\pi k\frac{m}{(N/2)}} \tag{22}$$

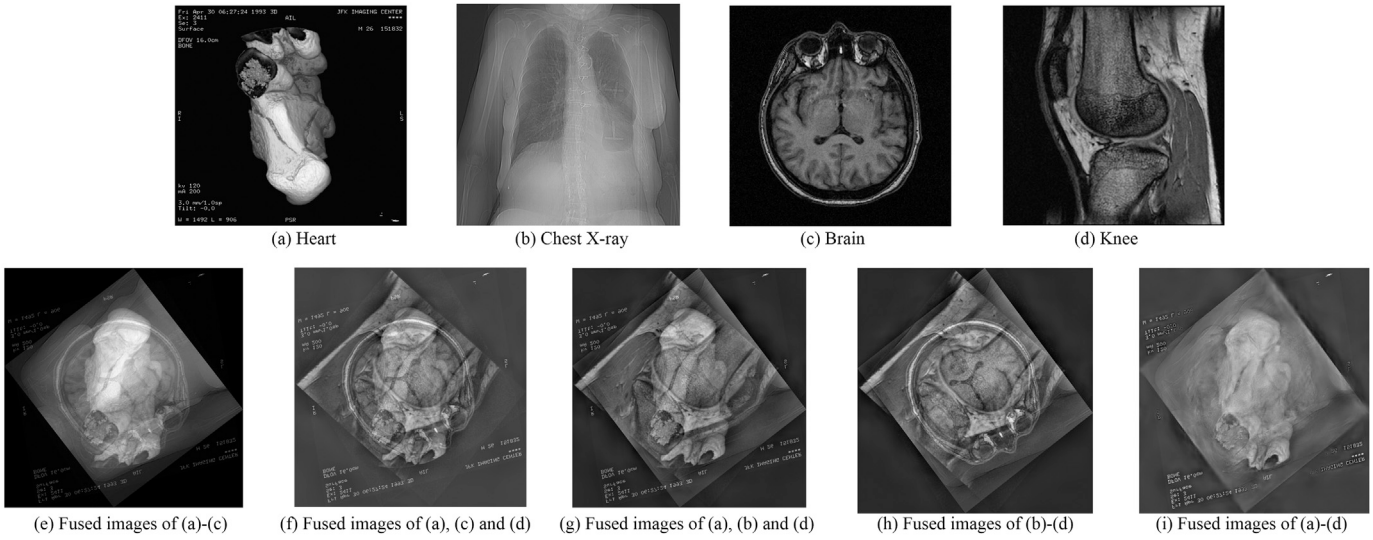
where $\sum_{m=0}^{N/2-1} x_{2m} \cdot e^{-i2\pi k\frac{m}{N/2}}$ represents the even-indexed DFT and $e^{-i2\pi k\frac{1}{N}}\sum_{m=0}^{N/2-1} x_{2m+1} \cdot e^{-i2\pi k\frac{m}{N/2}}$ means the odd-indexed DFT.

3. Effect of the attacks on the stability of extracted feature vectors

The performance of zero-watermarking methods against attacks mainly depends on whether the essential features extracted when constructing a zero-watermarking signal exhibit strong robustness against attacks. In this study, we first normalized and fused multiple images. Then, we extracted the effective regions of the fused images and performed CT and CHFMs to generate the magnitude sequence. Finally, an FFT was performed on the generated magnitude sequence to obtain 64-bit feature vectors. To validate the ability of the proposed method to resist attacks, the following two experiments were conducted:

(1) The stability of the extracted feature vectors of the cover image against various attacks was verified on the Chest X-ray image shown in Fig. 5. Table 1 shows the corresponding results. As observed, the extracted feature vectors (64 bits) under different attacks are almost unchanged, and the correlation coefficients are all higher than 0.984, indicating that the extracted feature vectors exhibit strong robustness in the face of various attacks.

(2) The uniqueness of the feature vectors generated from the fused images was verified on the feature vectors extracted from the images in Fig. 5 after fusion. The experimental results are shown in Tables 2 and 3, where P_1, P_2, P_3, and P_4 denote the Heart, Chest X-ray, Brain, and Knee images, respectively. The results show that the extracted feature vectors from different fused images differ, with a similarity of approximately 0.5. In contrast, the feature vectors from the same



(a) Heart (b) Chest X-ray (c) Brain (d) Knee

(e) Fused images of (a)-(c) (f) Fused images of (a), (c) and (d) (g) Fused images of (a), (b) and (d) (h) Fused images of (b)-(d) (i) Fused images of (a)-(d)

Fig. 5. Four original medical images and their fusion.

Table 1
Feature vectors generated under different attacks (64-bit).

Type of attack	Generated feature vectors	NC
No attacks	1111111100111111000000000000001100000000000000001011111111111111	–
JPEG compression (QF = 15)	1111111100111111000000000000001100000000000000011111111111111111	0.998
Median filtering (3 × 3)	1111111100111111000000000000001100000000000000001111111111111111	0.999
Wiener filtering (3 × 3)	1111111100111111000000000000001100000000000000001011111111111111	1.000
Gaussian noise (0.1)	1111111101111111000000000000001100000000000000011111111111111111	0.991
Salt & pepper noise (0.1)	1111111101111111000000000000001100000000000000011111111111111111	0.993
Rotation attack (10°)	1111111101101111000000000000001100000000000000001011111111111111	0.984
Scaling attack (Shrink 0.25x)	1111111100111111000000000000001100000000000000001011111111111111	1.000
Cropping attack (Upper left 1/16)	1111111110011111000000000000001100000000000000001011111111111111	0.992

Table 2
Feature vectors generated by different images fusion (64-bit).

Fusion of different images	Generated feature vectors
P_1, P_2, P_3	0000001110000111111101111110111111100111110001111000000000001111
P_1, P_2, P_4	0110001111111111111111111111111111111111111111111111111000000000000011
P_2, P_3, P_4	0000111011100111000011111000111000111111000000000000000000000000
P_1, P_3, P_4	0001111111111111111101111111111111111111111111111111000000000000111
P_1, P_2, P_3, P_4	0001000000000011101111100011111111111111111110001111111111111000111

Table 3
Similarity of feature vectors generated from different images fusion.

Fusion of different images	P_1, P_2, P_3	P_1, P_2, P_4	P_2, P_3, P_4	P_1, P_3, P_4	P_1, P_2, P_3, P_4
P_1, P_2, P_3	1.000	0.574	0.554	0.546	0.528
P_1, P_2, P_4	0.576	1.000	0.501	0.512	0.563
P_2, P_3, P_4	0.552	0.501	1.000	0.581	0.515
P_1, P_3, P_4	0.546	0.512	0.591	1.000	0.530
P_1, P_2, P_3, P_4	0.528	0.563	0.515	0.530	1.000

fused images are identical, with a similarity of 1.000. This indicates that the extracted feature vectors can effectively distinguish the fusion of different images.

The experimental results demonstrate that the constructed feature signal exhibits robust performance, providing a theoretical basis for utilizing the feature signal to generate a robust zero-watermarking signal.

4. Proposed method

To address the poor performance of most methods in resisting diversity attacks and the high storage space required for centralized protection of multiple images, a robust zero-watermarking method combining image moments and multi-scale transformation is proposed. Figs. 6–8 show the flowcharts for the copyrighted image encryption and decryption, zero-watermarking construction, and detection algorithms, respectively.

4.1. Copyrighted image encryption

To enhance the security of the method, a copyrighted image CI of size $m \times n$ was encrypted using the Lorenz chaotic system and Fibonacci Q-matrix. Fig. 6 shows the experimental results after encrypting the copyrighted image using the following key steps:

Step 1: Using the original copyrighted image CI of size $m \times n$, the initial key x_1 of the Lorenz chaotic system is computed.

$$x_1 = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} CI(i,j) + (m \times n)}{2000 + (m \times n)} \tag{23}$$

Two new values, x_2 and x_3, are then obtained by iterating twice. Finally, x_1, x_2, and x_3 are chosen as the initial values of the state variables x, y, and z, respectively.

Step 2: Based on the selected initial values, three vectors, X, Y and Z are generated using Eq. (1), from which three sub-vectors of length $m \times \frac{n}{3}$ are chosen to construct a vector V of length $m \times n$.

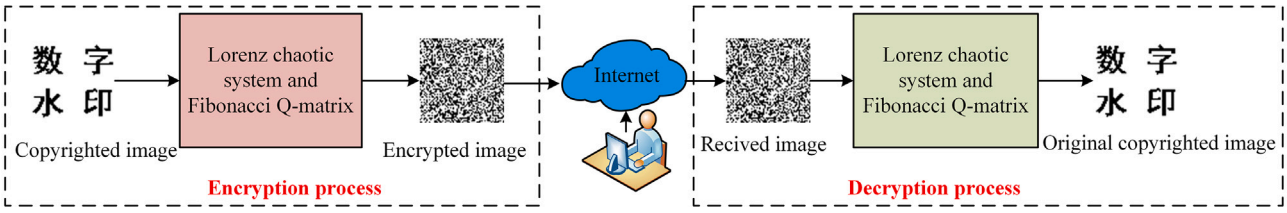


Fig. 6. An example of simple copyrighted image encryption and decryption.

Step 3: The copyrighted image CI is first reshaped into a one-dimensional vector G, and then, the sequence V is sorted in ascending order to obtain index IX. Finally, G is permuted using IX to generate a scrambled one-dimensional vector R.

Step 4: The R vector is reshaped into a matrix of size $m \times n$, and the matrix is partitioned into blocks of size 2×2.

Step 5: Set the parameter $n = 20$ in Eq. (5) to compute Q^n. Then, perform a modulo-2 operation on each term in Q^n to obtain a binary matrix.

Step 6: Using the Fibonacci Q-matrix construction method introduced in Section 2.2, an exclusive-OR operation is performed between each block of size 2×2 and the Fibonacci Q-matrix to obtain an encrypted copyrighted image (ECI).

The image decryption step is simply the reverse of the encryption step and is not described here.

4.2. Zero-watermarking construction

Assuming that the sizes of the four cover images I and the copyrighted image CI are $M \times N$ and $m \times n$, respectively. A robust feature image is constructed by combining image moments and multi-scale transforms, and a robust zero-watermarking signal is generated by performing an exclusive-OR (XOR) operation with the encrypted copyrighted image. The key steps of the proposed method are outlined as follows.

Step 1: Using the moment-based image normalization technique in Section 2.3, four gray-scale images of size $M \times N$ are subjected to the corresponding normalization process. Then, scaling and rotation normalizations are applied to obtain four standard normalized images.

Step 2: A new fused image (FI) is generated by fusing the information of the four normalized images using the image fusion method in Section 2.4.

Step 3: For a fused image FI of size $M \times N$, the geometric center of FI is defined as $x = \frac{M}{2}$, $y = \frac{N}{2}$. The effective region (ER) of size $P \times Q$ is extracted from the fused image FI using Eq. (24).

$$ER = \mathrm{FI}\left[(x - \frac{P}{2}) : (x + \frac{P}{2} - 1), (y - \frac{Q}{2}) : (y + \frac{Q}{2} - 1)\right] \tag{24}$$

Step 4: Using the Contourlet transform, the LSs are obtained from the extracted ER. A square region (SR) of size $((M+N)/2) \times ((M+N)/2)$ is then selected from LSs.

Step 5: The maximum-order $n_{\max} = 25$ is selected, and the region SR is computed using Eq. (15) to obtain $(n_{\max}+1)(2n_{\max}-1)$ CHFMs.

Step 6: To make the number of CHFMs the same size as the copyrighted image, $m \times n$ moment values are obtained by expanding the amplitude sequence of the $(n_{\max}+1)(2n_{\max}-1)$ moments, converting them into an $m \times n$ one-dimensional vector $A = \{a(i), 1 \le i \le m \times n\}$.

Step 7: FFT is performed on one-dimensional vector A to generate one-dimensional vector $B = \{b(i), 1 \le i \le m \times n\}$.

Step 8: Reshape the vector B into a two-dimensional matrix C. Calculate the mean value M of the matrix C and binarize it using M as a threshold. Specifically, if the value of an element of C is greater than or equal to M, the feature bit is 1; otherwise, the feature bit is 0. This method is used to construct the binary feature image $F = \{f(i,j), 1 \le i \le m, 1 \le j \le n\}$.

$$F(i,j) = \begin{cases} 1, C(i,j) \ge M \\ 0, C(i,j) < M \end{cases} \tag{25}$$

Step 9: Perform an XOR operation between the feature matrix F obtained in Step 8 and the encrypted copyrighted image ECI in Section 4.1 to get a robust zero-watermarking image, which is then authenticated and registered with a third-party intellectual property rights (IPR). The unique ID number is then saved as the basis for copyright extraction. The zero-watermarking image construction and registration processes are thus completed.

$$Z = \mathrm{XOR}(ECI, F) \tag{26}$$

4.3. Zero-watermarking detection

The zero-watermarking detection process is the reverse of the zero-watermarking construction method. Below is a description of the key steps.

Step 1: Same as Step 1 of the zero-watermarking signal generation process, four gray-scale images of size $M \times N$ are normalized using the method described in Section 2.3, followed by scaling and rotation normalizations to produce standard normalized images.

Step 2: The corresponding feature image is obtained by performing the normalized standard images following Steps 2–8 in Section 4.2.

Step 3: A zero-watermarking image saved by a third-party authentication center can be obtained using the ID number. Then, an XOR operation is performed on the zero-watermarking image and the generated feature image, resulting in an undecrypted copyright image (UCI).

$$UCI = \mathrm{XOR}(Z, F') \tag{27}$$

Step 4: The original copyrighted image CI can be recovered by decrypting the undecrypted copyrighted image UCI using the Lorenz chaotic system and the Fibonacci Q-matrix. Because the original CI is a meaningful and recognizable image, the human eye can directly authenticate the recovered copyrighted image.

5. Experimental results and analysis

5.1. Experimental parameters

To verify the effectiveness of our method, a simulation experiment was conducted in two software environments: one configured with MATLAB R2023a and the other with Microsoft Windows 11. Four 512×512 standard medical images: Heart, Chest X-ray, Brain, and Knee were chosen as experimental images, as shown in Fig. 9(a)~(d). Fig. 9(e) shows the original binary copyrighted image, which is a 64×64 pixel binary image composed of a binary sequence of length 4096. Fig. 9(f) displays a zero-watermarking image generated by this proposed method. As can be seen, the resulting zero-watermarking image looks cluttered and, if not recovered, unrecognizable to the human eye.

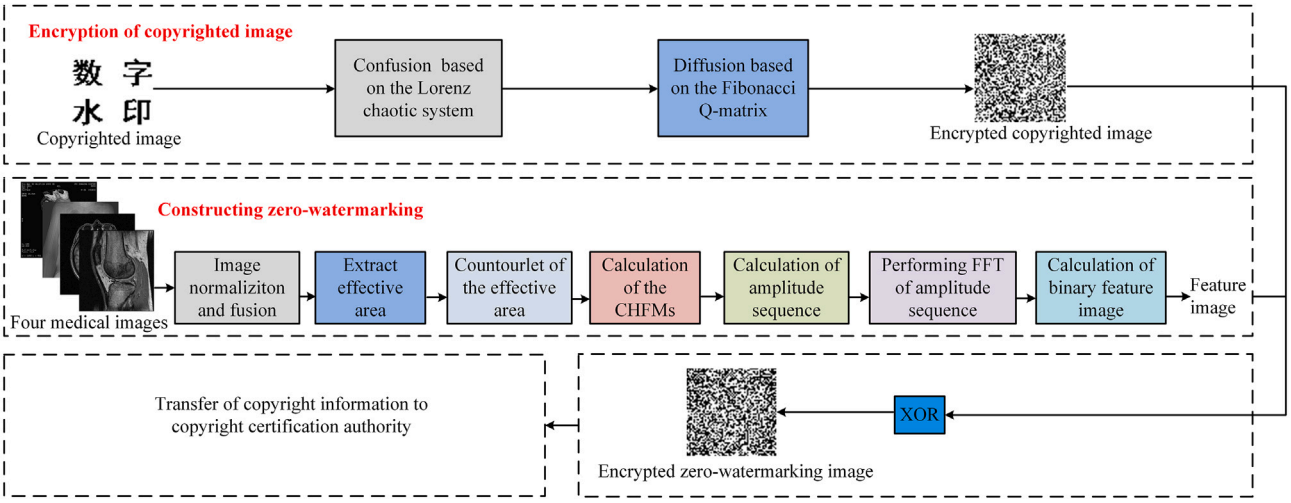


Fig. 7. Flowchart of zero-watermarking construction method.

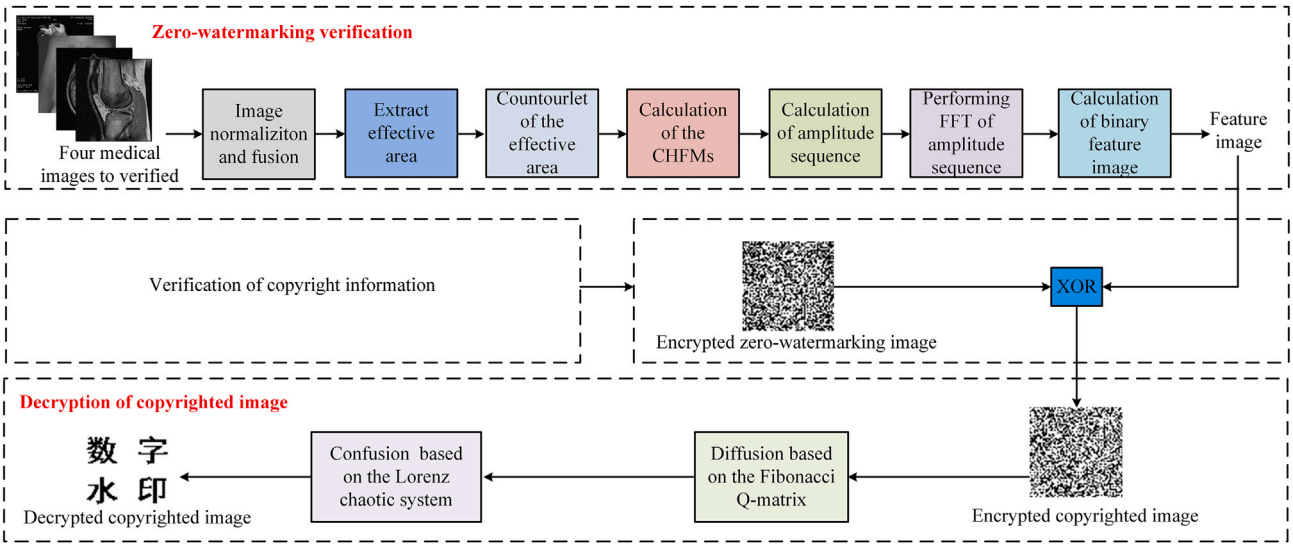


Fig. 8. Flowchart of zero-watermarking detection method.

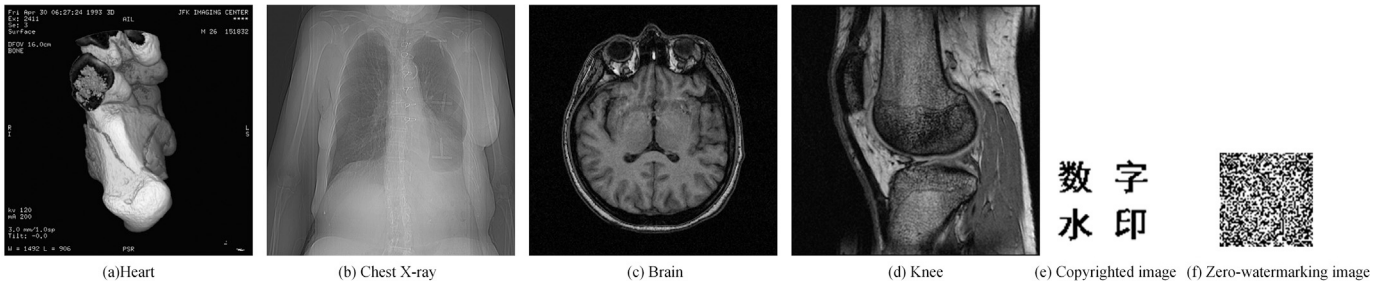


Fig. 9. Original medical image, original copyrighted image, and generated zero-watermarking image.

5.2. Evaluation indicators

The attack resistance of the methods is measured using a generalized normalized correlation coefficient (NC), and the quality of the reconstructed image is evaluated using a generalized mean-square reconstruction error (MSRE) to objectively assess the methods' performance.

(1) Normalized correlation

The NC value is commonly used to measure the similarity between a copyrighted image extracted from an attacked cover image and the original copyrighted image. The NC value typically falls between 0 and

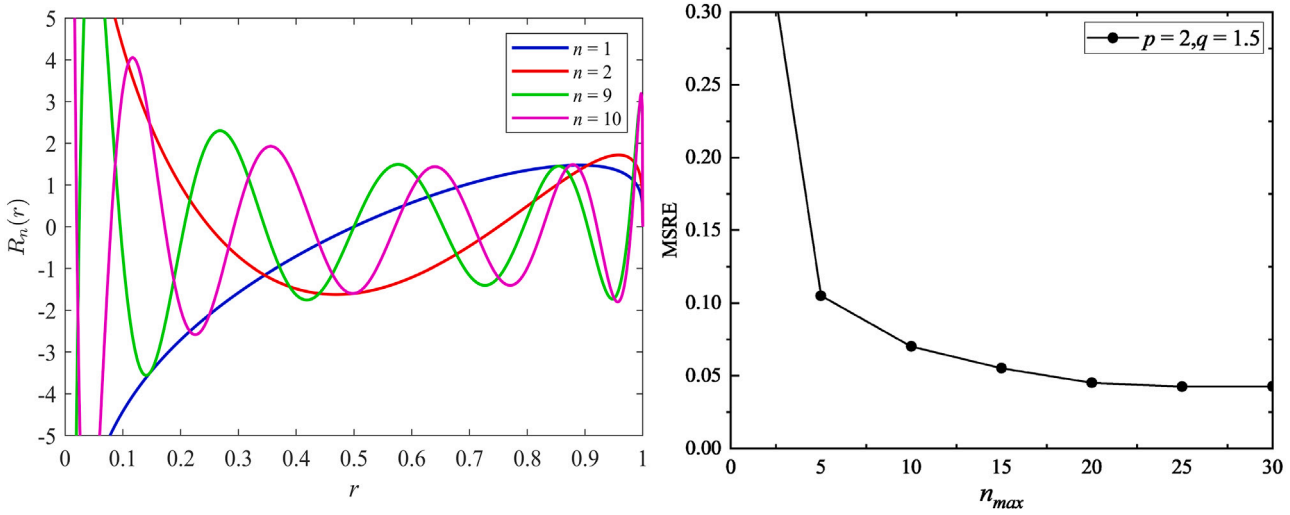


Fig. 10. (a) Variations in the value of $R_n(r)$ with r, in the interval $0 < r \leq 1$, $n_{\max} = 1, 2, 9, 10$. (b) MSRE values corresponding to CHFMs with different orders for the grayscale image Heart.

1, where 0 indicates that the two images are not similar, and 1 indicates that they are identical. In other words, the higher the NC value, the more similar the two images are, suggesting that the method is more resistant to attacks.

$$NC(OCI, ECI) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}[OCI(i,j)ECI(i,j)]}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}OCI(i,j)^2}\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}ECI(i,j)^2}} \quad (28)$$

where both *OCI* and *ECI* are of size $m \times n$; *OCI* refers to the original copyrighted image, while *ECI* is the copyrighted image extracted after the cover image has undergone an attack.

(2) Mean-squared reconstruction error

As a generalized tool, the quality of the reconstructed images can be objectively assessed using the MSRE in Eq. (29). In general, the smaller the MSRE value, the lower the error between the reconstructed and original images, indicating better image quality; conversely, a higher MSRE value suggests poorer reconstruction quality.

$$MSRE = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[I'(x,y) - I(x,y)\right]^2 dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[I'(x,y)]^2 dxdy} \quad (29)$$

where I and I' denote the original and reconstructed images, respectively.

5.3. Image reconstruction experiments

Fig. 10(a) shows the variation in the values of the radial polynomial function $R_n(r)$ (Eq. (16)) in the interval [0, 1]. It can be seen that $R_n(r)$ has n zeros, which satisfy a uniform distribution in the interval [0, 1], and the values of the function located near the zeros of different orders are almost the same.

To verify the reconstruction ability of the CHFMs, experiments were conducted by setting the parameters $p = 2$ and $q = 1.5$ and selecting a standard medical heart image of size 512×512. Figs. 10(b) and 11 show the corresponding MSRE values and reconstructed images for $n = 0, 5, \ldots, 25$. As shown in Fig. 10(b), the MSRE is the lowest when $n_{max} = 25$. The best-quality reconstructed image is observed in Fig. 11 for $n_{max} = 25$.

5.4. Resistance to regular attack experiments

In this section, the NC value is used to quantitatively assess the quality of the extracted copyrighted image, which reflects the method's resistance to attacks. The results show that when the cover image is not attacked, the NC value between the extracted and original copyrighted images is 1.000. In addition, the proposed method exhibited strong robustness when the cover image was attacked. To perform a systematic and robust assessment of the proposed method, images, as well as two-, three-, and four-fused images, were tested for their resistance to attacks. The detailed experiments are described below.

Table 4
Results of resisting JPEG compression.

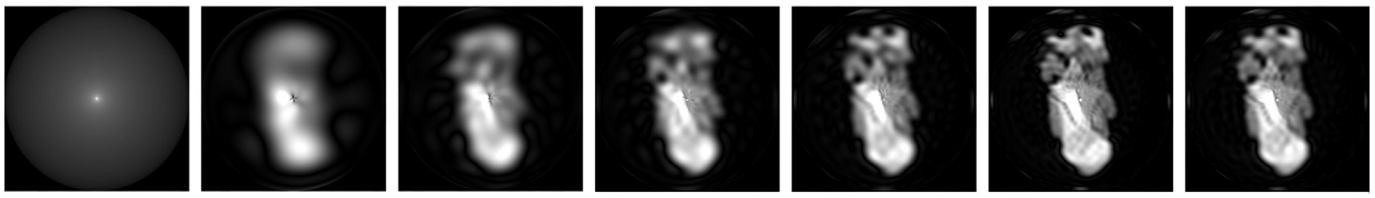
Fusion of different images	Quality factors (QF) 5	10	15	20	25	Average
P_1	0.985	0.990	0.986	0.989	0.996	0.989
P_2	0.996	0.999	0.998	1.000	0.999	0.998
P_3	0.995	0.994	0.997	0.998	0.998	0.996
P_4	0.987	0.996	0.998	1.000	0.999	0.996
P_1, P_2	0.992	0.994	0.998	0.998	0.999	0.996
P_1, P_3	0.991	0.991	0.993	0.993	0.993	0.992
P_1, P_4	0.989	0.995	0.994	0.994	0.995	0.993
P_2, P_3	0.994	0.996	0.996	0.998	0.998	0.996
P_2, P_4	0.994	0.997	0.998	0.999	0.999	0.997
P_3, P_4	0.989	0.996	0.998	1.000	0.999	0.996
P_1, P_2, P_3	0.991	0.997	0.998	0.998	0.999	0.996
P_1, P_2, P_4	0.994	0.996	0.998	0.998	0.999	0.997
P_2, P_3, P_4	0.994	0.995	0.998	0.997	0.999	0.996
P_1, P_3, P_4	0.987	0.991	0.989	0.990	0.989	0.989
P_1, P_2, P_3, P_4	0.994	0.999	0.999	0.998	0.998	0.997

5.4.1. JPEG compression attack

The ability to resist JPEG compression attacks is summarized in Table 4. It can be seen that the proposed method is more resistant to JPEG compression, with an average NC value of 0.995. This may be because the proposed method chooses to compute the CHFMs in the LSs of the CT transform, where the information is more concentrated, thereby enhancing its resistance to JPEG compression.

5.4.2. Noise attack

Table 5 summarizes the experimental results against Gaussian white noise and salt & pepper noise attacks, and Table 6 lists the results for Gaussian noise and speckle noise attacks. Note that for the Gaussian white noise attack, the values of the parameter *intensity* are 0, 0.5, and 1. It is observed that our method has high resistance to noise attacks with NC values of 0.968, 0.963, 0.958, and 0.989 against Gaussian white noise, salt & pepper noise, Gaussian noise, and speckle noise attacks, respectively. This may be because the technique used in the proposed method has a suppression effect on noise in the transform domain, which enhances its ability to resist noise attacks.



(a) $n_{max} = 0$ (b) $n_{max} = 5$ (c) $n_{max} = 10$ (d) $n_{max} = 15$ (e) $n_{max} = 20$ (f) $n_{max} = 25$ (g) $n_{max} = 30$

Fig. 11. Samples of CHFMs reconstructed with different orders.

Table 5
Results of resisting Gaussian white noise and salt & pepper noise attacks.

Fusion of different images	Gaussian white noise					Average	Salt & pepper noise					Average
	(0.05, 0.025,0)	(0.1, 0.05,0)	(0.2, 0.1,0)	(0.1, 0.1,0.1)	(0.5, 0.25,0)		0.1	0.2	0.3	0.4	0.5	
P_1	0.936	0.914	0.904	0.915	0.900	0.914	0.951	0.922	0.903	0.894	0.896	0.913
P_2	0.996	0.993	0.990	0.990	0.979	0.990	0.993	0.990	0.985	0.979	0.974	0.984
P_3	0.964	0.956	0.944	0.958	0.931	0.950	0.978	0.958	0.951	0.941	0.937	0.953
P_4	0.976	0.959	0.943	0.960	0.930	0.954	0.982	0.966	0.949	0.938	0.932	0.954
P_1, P_2	0.992	0.988	0.979	0.981	0.962	0.981	0.984	0.980	0.971	0.959	0.950	0.969
P_1, P_3	0.978	0.963	0.956	0.963	0.937	0.959	0.978	0.968	0.950	0.940	0.925	0.952
P_1, P_4	0.979	0.973	0.959	0.970	0.948	0.966	0.980	0.971	0.953	0.950	0.935	0.958
P_2, P_3	0.994	0.991	0.985	0.986	0.976	0.986	0.990	0.984	0.980	0.979	0.973	0.981
P_2, P_4	0.994	0.988	0.979	0.985	0.966	0.982	0.991	0.983	0.976	0.971	0.958	0.976
P_3, P_4	0.984	0.970	0.963	0.974	0.948	0.968	0.986	0.974	0.964	0.956	0.952	0.966
P_1, P_2, P_3	0.991	0.980	0.974	0.976	0.957	0.976	0.987	0.971	0.966	0.959	0.953	0.967
P_1, P_2, P_4	0.988	0.979	0.971	0.977	0.963	0.976	0.986	0.978	0.968	0.963	0.956	0.970
P_2, P_3, P_4	0.987	0.982	0.974	0.980	0.961	0.977	0.986	0.981	0.973	0.961	0.953	0.971
P_1, P_3, P_4	0.978	0.971	0.961	0.967	0.947	0.965	0.978	0.969	0.957	0.949	0.936	0.958
P_1, P_2, P_3, P_4	0.990	0.984	0.976	0.972	0.965	0.977	0.997	0.994	0.989	0.990	0.987	0.991

Table 6
Results of resisting Gaussian noise and speckle noise attacks.

Fusion of different images	Gaussian noise					Average	Speckle noise					Average
	0.1	0.2	0.3	0.4	0.5		0.1	0.2	0.3	0.4	0.5	
P_1	0.926	0.903	0.894	0.898	0.894	0.903	0.993	0.987	0.983	0.983	0.977	0.985
P_2	0.991	0.986	0.977	0.972	0.960	0.977	0.996	0.995	0.993	0.992	0.992	0.993
P_3	0.965	0.954	0.947	0.939	0.932	0.947	0.997	0.996	0.994	0.994	0.992	0.995
P_4	0.971	0.952	0.939	0.934	0.930	0.945	0.997	0.997	0.993	0.991	0.990	0.993
P_1, P_2	0.984	0.971	0.959	0.950	0.944	0.962	0.991	0.988	0.989	0.979	0.978	0.985
P_1, P_3	0.970	0.957	0.941	0.931	0.919	0.944	0.992	0.986	0.987	0.980	0.981	0.986
P_1, P_4	0.977	0.963	0.949	0.938	0.933	0.952	0.988	0.988	0.987	0.980	0.980	0.985
P_2, P_3	0.989	0.984	0.980	0.977	0.974	0.981	0.996	0.993	0.992	0.991	0.989	0.992
P_2, P_4	0.987	0.979	0.970	0.962	0.957	0.971	0.995	0.994	0.986	0.987	0.987	0.990
P_3, P_4	0.976	0.970	0.960	0.956	0.950	0.963	0.997	0.995	0.989	0.989	0.990	0.992
P_1, P_2, P_3	0.983	0.967	0.961	0.950	0.942	0.961	0.992	0.988	0.985	0.984	0.979	0.986
P_1, P_2, P_4	0.979	0.970	0.964	0.957	0.951	0.964	0.989	0.992	0.989	0.983	0.984	0.988
P_2, P_3, P_4	0.982	0.975	0.966	0.959	0.956	0.968	0.994	0.996	0.985	0.990	0.986	0.990
P_1, P_2, P_4	0.973	0.960	0.952	0.940	0.935	0.952	0.990	0.991	0.985	0.983	0.986	0.987
P_1, P_2, P_3, P_4	0.990	0.984	0.976	0.972	0.965	0.977	0.997	0.994	0.989	0.990	0.987	0.991

5.4.3. Filtering attack

Table 7 lists the experimental results for Median and Wiener filtering attacks, and Table 8 gives the experimental results for Gaussian low-pass and mean filtering attacks. As observed, our method has high resistance to filtering attacks with NC values of 0.994, 0.997, 1.000, and 0.994 against median filtering, Wiener filtering, Gaussian low-pass filtering, and mean filtering attacks, respectively. This may be because the CT transform used in the proposed method provides a nuanced and compelling characterization of the local and global features of the cover image in the transform domain, which makes the cover image highly stable when subjected to a filtering attack, effectively enhancing its ability to resist the filtering attack.

5.5. Resistance to geometric attack experiments

5.5.1. Offset rows, columns, and cropping attacks

Table 9 provides experimental results for offset rows, columns, and cropping attacks. The proposed method exhibits robustness against offset rank attacks and cropping attacks, with an average NC of 0.989 against offset rank attacks and 0.965 against cropping attacks. This finding can be attributed to two key reasons. First, the orthogonality of the Chebyshev polynomials is independent of each other within a specific interval, which helps reduce interference between different frequency components, thereby improving the stability and robustness of the signal. Second, the FFT converts the amplitude signal from the time domain to the frequency domain, resulting in the loss of key information when the original signal is affected by an offset or cropping attack in the time domain. In contrast, the information remains relatively intact in the frequency domain. Consequently, the proposed method can effectively resist offset-rank and cropping attacks.

5.5.2. Scaling attack

The results of the scaling attack are summarized in Table 10. Note that after scaling the image by a factor of x, it needs to be scaled again by a factor of $\frac{1}{x}$ before constructing the feature image. As seen, the proposed method demonstrated outstanding resistance to scaling

Table 7
Results of resisting median and Wiener filtering attacks.

Fusion of	Median filtering					Average	Wiener filtering					Average
different images	3 × 3	5 × 5	7 × 7	9 × 9	11 × 11		3 × 3	5 × 5	7 × 7	9 × 9	11 × 11	
P_1	0.990	0.973	0.968	0.966	0.965	0.972	0.998	0.998	0.998	0.998	0.998	0.998
P_2	0.999	0.999	0.999	0.999	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
P_3	1.000	0.997	0.995	0.994	0.992	0.995	1.000	0.999	0.998	0.997	0.997	0.998
P_4	1.000	0.999	0.997	0.996	0.995	0.997	1.000	1.000	1.000	1.000	0.999	1.000
P_1, P_2	0.998	0.996	0.996	0.995	0.993	0.996	0.999	0.998	0.997	0.997	0.995	0.997
P_1, P_3	1.000	0.997	0.994	0.992	0.988	0.994	1.000	0.999	0.998	0.994	0.992	0.996
P_1, P_4	0.999	0.996	0.994	0.990	0.986	0.993	1.000	0.998	0.995	0.991	0.989	0.994
P_2, P_3	0.999	0.999	0.998	0.995	0.993	0.997	1.000	0.999	0.998	0.997	0.995	0.998
P_2, P_4	0.998	0.997	0.994	0.994	0.993	0.995	0.999	0.998	0.994	0.993	0.992	0.995
P_3, P_4	0.999	0.997	0.994	0.991	0.986	0.993	0.998	0.998	0.997	0.996	0.995	0.997
P_1, P_2, P_3	0.999	0.997	0.995	0.994	0.990	0.995	0.999	0.998	0.995	0.995	0.995	0.996
P_1, P_2, P_4	0.998	0.994	0.990	0.989	0.987	0.992	0.998	0.995	0.992	0.989	0.987	0.992
P_2, P_3, P_4	0.999	0.996	0.993	0.990	0.985	0.993	0.999	0.997	0.993	0.991	0.989	0.994
P_1, P_3, P_4	0.998	0.998	0.996	0.992	0.989	0.995	0.999	0.998	0.998	0.996	0.995	0.997
P_1, P_2, P_3, P_4	1.000	0.999	0.998	0.998	0.997	0.998	1.000	1.000	0.999	0.998	0.997	0.999

Table 8
Results of resisting Gaussian low-pass and mean filtering attacks.

Fusion of	Gaussian low-pass filtering					Average	Mean filtering					Average
different images	3 × 3	5 × 5	7 × 7	9 × 9	11 × 11		3 × 3	5 × 5	7 × 7	9 × 9	11 × 11	
P_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997	0.996	0.996	0.997
P_2	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.998	0.998	0.998	0.997	0.998
P_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.996	0.996	0.995	0.997
P_4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.998	0.997	0.998
P_1, P_2	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.995	0.994	0.992	0.990	0.993
P_1, P_3	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.996	0.994	0.990	0.990	0.993
P_1, P_4	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.994	0.991	0.990	0.984	0.991
P_2, P_3	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.995	0.993	0.991	0.995
P_2, P_4	0.999	0.999	0.999	0.999	0.999	0.999	0.995	0.993	0.991	0.991	0.989	0.992
P_3, P_4	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.995	0.993	0.992	0.988	0.993
P_1, P_2, P_3	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.994	0.993	0.991	0.988	0.992
P_1, P_2, P_4	0.999	0.999	0.999	0.999	0.999	0.999	0.996	0.993	0.988	0.986	0.986	0.990
P_2, P_3, P_4	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.994	0.989	0.986	0.985	0.990
P_1, P_3, P_4	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.996	0.995	0.991	0.989	0.994
P_1, P_2, P_3, P_4	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.997	0.995	0.994	0.992	0.995

Table 9
Results of resisting offset and cropping attacks.

Fusion of	Offset direction				Average	Cropping position				Average
different images	Shift right 2 columns	Shift left 2 columns	Shift up 2 rows	Shift down 2 rows		Upper left 1/16	Upper left 1/8	Upper left 1/4	Center 1/4	
P_1	0.995	0.993	0.992	0.994	0.993	0.976	0.963	0.923	0.933	0.949
P_2	0.997	0.995	0.995	0.997	0.996	0.992	0.987	0.962	0.979	0.980
P_3	0.996	0.992	0.993	0.993	0.993	0.983	0.973	0.922	0.955	0.958
P_4	0.993	0.995	0.998	0.997	0.996	0.979	0.963	0.932	0.952	0.957
P_1, P_2	0.991	0.992	0.985	0.986	0.989	0.993	0.981	0.958	0.982	0.978
P_1, P_3	0.985	0.984	0.988	0.984	0.985	0.980	0.970	0.941	0.968	0.965
P_1, P_4	0.981	0.981	0.984	0.979	0.981	0.987	0.967	0.928	0.948	0.958
P_2, P_3	0.989	0.993	0.992	0.986	0.990	0.989	0.979	0.960	0.971	0.975
P_2, P_4	0.988	0.988	0.988	0.987	0.988	0.992	0.981	0.959	0.956	0.972
P_3, P_4	0.986	0.989	0.988	0.982	0.986	0.987	0.962	0.925	0.930	0.951
P_1, P_2, P_3	0.989	0.989	0.988	0.985	0.988	0.991	0.981	0.952	0.975	0.975
P_1, P_2, P_4	0.986	0.987	0.985	0.982	0.985	0.991	0.974	0.945	0.964	0.969
P_2, P_3, P_4	0.987	0.987	0.987	0.981	0.986	0.989	0.974	0.948	0.951	0.966
P_1, P_3, P_4	0.986	0.982	0.986	0.979	0.983	0.979	0.960	0.924	0.953	0.954
P_1, P_2, P_3, P_4	0.992	0.995	0.992	0.989	0.992	0.993	0.983	0.958	0.963	0.974

attacks, as evidenced by its average NC value of 0.998. The main reasons for this are as follows: The CT transform can effectively capture the local features of the cover image, and the method's resistance to scaling attacks is improved by the normalizing process based on image moments. These properties enable the proposed method to maintain the stability of the extracted feature vectors when an image undergoes a scaling attack. Consequently, the proposed method has enhanced its ability to resist scaling attacks.

5.5.3. Rotation attack

The results of rotation attacks are summarized in Table 11. It can be seen that the proposed method achieves strong resistance to rotation attack, with an average NC value of 0.964. It is mainly attributed to the fact that CHFMs possess rotational invariance when computing the LSs. This property ensures that even if the LSs are rotated, the feature data can still be effectively extracted in the low-frequency part. Additionally, the FFT transforms the amplitude sequence, further enhancing the

Table 10
Results of resisting scaling attack.

Fusion of different images	Scaling								Average
	Shrink 0.25x	Shrink 0.4x	Shrink 0.5x	Shrink 0.7x	Magnify 4x	Magnify 2x	Magnify 2.4x	Magnify 1.3x	
P_1	0.999	0.999	1.000	0.998	1.000	1.000	1.000	0.999	0.999
P_2	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000
P_3	0.998	0.998	1.000	0.997	1.000	1.000	1.000	0.999	0.999
P_4	0.999	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000
P_1, P_2	0.998	0.993	1.000	0.994	1.000	1.000	0.998	0.999	0.998
P_1, P_3	0.997	0.994	1.000	0.995	1.000	1.000	0.998	0.999	0.998
P_1, P_4	0.996	0.988	0.999	0.992	1.000	1.000	0.999	0.999	0.997
P_2, P_3	0.999	0.996	0.999	0.996	1.000	1.000	1.000	1.000	0.998
P_2, P_4	0.997	0.992	0.999	0.994	1.000	1.000	0.999	0.999	0.997
P_3, P_4	0.998	0.993	1.000	0.994	1.000	1.000	1.000	1.000	0.998
P_1, P_2, P_3	0.998	0.992	1.000	0.995	1.000	1.000	0.999	0.999	0.998
P_1, P_2, P_4	0.994	0.989	0.999	0.992	1.000	1.000	0.997	0.999	0.996
P_2, P_3, P_4	0.996	0.989	1.000	0.990	1.000	1.000	0.998	0.999	0.996
P_1, P_3, P_4	0.996	0.991	0.999	0.992	1.000	1.000	0.998	0.999	0.997
P_1, P_2, P_3, P_4	1.000	0.999	1.000	0.999	1.000	1.000	1.000	1.000	1.000

Table 11
Results of resisting rotation attack.

Fusion of different images	Rotation angle									Average
	10°	20°	30°	40°	50°	60°	70°	80°	90°	
P_1	0.973	0.958	0.956	0.955	0.956	0.960	0.965	0.980	0.996	0.967
P_2	0.984	0.970	0.958	0.940	0.941	0.958	0.972	0.984	0.999	0.967
P_3	0.986	0.981	0.982	0.979	0.981	0.982	0.987	0.988	0.995	0.985
P_4	0.975	0.951	0.937	0.924	0.924	0.941	0.961	0.979	0.997	0.954
P_1, P_2	0.998	1.000	1.000	1.000	0.993	0.994	0.998	0.999	0.993	0.997
P_1, P_3	0.981	0.969	0.967	0.962	0.958	0.963	0.961	0.971	0.992	0.969
P_1, P_4	0.973	0.965	0.962	0.950	0.938	0.947	0.954	0.973	0.994	0.962
P_2, P_3	0.980	0.959	0.950	0.934	0.935	0.949	0.958	0.975	0.998	0.960
P_2, P_4	0.969	0.948	0.925	0.912	0.899	0.925	0.954	0.969	0.995	0.944
P_3, P_4	0.978	0.957	0.937	0.922	0.916	0.939	0.952	0.974	0.996	0.952
P_1, P_2, P_3	0.978	0.962	0.957	0.947	0.942	0.944	0.956	0.977	0.994	0.962
P_1, P_2, P_4	0.969	0.968	0.954	0.944	0.933	0.937	0.949	0.971	0.991	0.957
P_2, P_3, P_4	0.972	0.948	0.933	0.920	0.918	0.936	0.949	0.973	0.994	0.949
P_1, P_3, P_4	0.978	0.964	0.961	0.955	0.950	0.951	0.958	0.979	0.988	0.965
P_1, P_2, P_3, P_4	0.978	0.969	0.956	0.944	0.938	0.950	0.959	0.979	0.999	0.963

rotational invariance in the frequency domain and thereby increasing the robustness of the method against rotation attacks. Consequently, the proposed method improves the resistance to rotation attacks.

5.6. Combined attack

To further measure the anti-attack capability of the method, the cover image was subjected to combined attacks, and the corresponding results are listed in Table 12. As shown in the table, the average NC value of the suggested method against the combined attacks is still as high as 0.980. According to the results above, the proposed method is capable of resisting a range of combined attacks, in addition to conventional and geometric attacks, indicating that it can withstand various types of attacks and exhibits strong, robust performance.

5.7. Impact of multiple images fusion on the performance of the proposed method

To objectively evaluate the impact of multiple image fusion on the performance of the proposed method, nine sets of images were first randomly selected from the image dataset BossBase [44], with the numbers of 5, 10, 15, 20, 30, 40, 60, 80, and 100, respectively. Then, for each set of images, a fused image is generated using the fusion technique described in Section 2.4. The proposed method is then utilized to construct a zero-watermarking image for the fused image and to perform experiments on various attacks. The experimental results are shown in Tables 13 and 14, from which it can be seen that the proposed method exhibits stable robustness under different numbers of image fusion conditions, with the average NC value always staying above 0.970, and the performance difference between various groups is not apparent. These results demonstrate that even when the number of fused images increases dramatically, the proposed method can still effectively resist multiple types of attacks and can be applied to the fusion needs of different numbers of images.

5.8. Experiments on image datasets

To verify the generalizability of the proposed method, we conduct experiments on four benchmark image datasets: BossBase [44], BOWS-2 [45], COVID [46], and SIPI [47]. For the experiments, 100 images were randomly selected from each image dataset for evaluation. The proposed method is first used to construct a zero-watermarking image for each test image. Then, an anti-attack test is performed to quantify the performance by calculating the NC value between the extracted image and the original copyrighted image. The average test results and the standard deviation STD of the 100 images are shown in Table 15. It can be seen that the average NC values of the proposed method are always higher than 0.95, and the STDs are less than 0.032 in all datasets, indicating that the proposed method not only exhibits excellent robustness on different datasets but also has excellent generalization ability. Although the experiments are conducted on standard datasets, the possible attacks on real-world natural images and medical images are simulated, which validates the ability of our method to resist attacks and generalization. COVID [46] is a publicly open dataset of chest X-rays and CT images of patients, containing 930 images. The proposed method demonstrates superior attack resistance on this dataset, indicating its potential application in real-world image copyright protection scenarios.

Table 12
Results of resisting combination attacks.

Fusion of different images	Type of combination attack							Average
	Rotation (10°) + Scaling (Shrink 0.25x)	JPEG (QF = 5) + Scaling (Shrink 0.25x)	JPEG (QF = 10) + Wiener filtering (3 × 3)	Gaussian noise (0.2) + Rotation (10°)	Gaussian noise (0.2) + Wiener filtering (3 × 3)	Median filtering (3 × 3) + Salt & pepper noise (0.2)	Median filtering (3 × 3) + Gaussian noise (0.2)	
P_1	0.979	0.987	0.989	0.903	0.902	0.920	0.903	0.943
P_2	0.967	0.997	0.998	0.986	0.985	0.990	0.985	0.991
P_3	0.957	0.995	0.995	0.954	0.955	0.958	0.953	0.972
P_4	0.979	0.989	0.995	0.953	0.953	0.967	0.953	0.973
P_1, P_2	0.973	0.992	0.996	0.970	0.970	0.978	0.970	0.981
P_1, P_3	0.967	0.991	0.992	0.958	0.956	0.968	0.957	0.974
P_1, P_4	0.956	0.987	0.994	0.960	0.964	0.973	0.962	0.977
P_2, P_3	0.978	0.996	0.998	0.984	0.985	0.987	0.986	0.990
P_2, P_4	0.975	0.993	0.999	0.982	0.980	0.983	0.978	0.987
P_3, P_4	0.967	0.988	0.998	0.969	0.970	0.973	0.971	0.981
P_1, P_2, P_3	0.977	0.992	0.997	0.968	0.968	0.977	0.970	0.981
P_1, P_2, P_4	0.963	0.994	0.994	0.970	0.969	0.975	0.970	0.981
P_2, P_3, P_4	0.964	0.994	0.995	0.978	0.976	0.980	0.975	0.985
P_1, P_3, P_4	0.976	0.985	0.991	0.959	0.960	0.963	0.960	0.974
P_1, P_2, P_3, P_4	0.978	0.995	0.999	0.982	0.983	0.983	0.983	0.989

Table 13
Experimental results of multi-image fusion against common attacks.

Number of fusion images	JPEG compression (QF = 15)	Median filtering (3 × 3)	Wiener filtering (3 × 3)	Gaussian low-pass filtering (3 × 3)	Mean filtering (3 × 3)	Gaussian noise (0.1)
5 images	0.998	1.000	1.000	1.000	0.994	0.994
10 images	0.998	1.000	1.000	1.000	0.998	0.989
15 images	1.000	1.000	1.000	1.000	0.999	0.991
20 images	0.997	0.999	0.999	1.000	0.996	0.989
30 images	0.996	1.000	1.000	1.000	0.996	0.990
40 images	0.999	1.000	1.000	1.000	0.996	0.990
60 images	0.997	0.999	0.999	1.000	0.996	0.980
80 images	0.998	1.000	1.000	1.000	0.990	0.979
100 images	0.999	1.000	1.000	1.000	0.998	0.992

Table 14
Experimental results of multi-image fusion against geometric attack.

Number of fusion images	Salt & pepper noise (0.1)	Speckle noise (0.1)	Gaussian white noise (0.1,0.05,0)	Rotation attack (10°)	Scaling attack (Shrink 0.25x)	Cropping attack (Upper left 1/16)
5 images	0.994	0.997	0.994	0.970	0.998	0.992
10 images	0.992	0.994	0.994	0.974	0.999	0.996
15 images	0.993	0.994	0.993	0.983	1.000	0.998
20 images	0.982	0.994	0.988	0.974	0.997	0.994
30 images	0.992	0.991	1.000	0.992	1.000	0.999
40 images	0.978	0.996	0.994	0.976	0.998	0.994
60 images	0.990	0.994	0.991	0.971	0.997	0.990
80 images	0.980	0.993	0.991	0.974	0.998	0.996
100 images	0.979	0.995	0.993	0.990	1.000	0.996

5.9. Comparison with similar methods

To highlight the superiority of the proposed method, six representative similar methods were selected for comparison experiments under the same conditions, and the results are shown in Table 16, where the proposed method is generally superior to the six similar methods in terms of robustness. The reasons for this can mainly be attributed to the following four aspects: First, the methods in [16–19] all use block processing, and the "block effect" introduced by these methods can lead to discontinuities or blurring of the boundaries between neighboring image blocks, which reduces the stability and accuracy of the feature vectors, whereas the proposed method generates the amplitude sequences by calculating the CHFMs of the effective regions of the LSs and performs the FFT transform. The proposed method not only avoids the "block effect" inherent in these methods but also leverages the rotational invariance and scaling invariance of CHFMs to construct feature vectors by computing CHFMs in the effective regions of the LSs. Performing the FFT transform further enhances the method's resistance to geometric attacks. Second, the methods in [22,25] both construct a zero-watermarking image based on image moments. The method in [22] employs FQGPCET, a nonlinear transformation method based on quaternions and polar coordinates, which is more sensitive to noise and shifts due to its nonlinear nature, potentially leading to distortion of the extracted features. The method in [25] is less robust to cropping and offset attacks due to the sensitivity of the polar harmonic invariant moments to cropping and offset. Specifically, the NC value obtained by the method is only 0.872 for the center 1/16 cropping attack since the cropping part is not used in the computation. However, the proposed method constructs binary eigenvectors using CHFMs and FFTs based on frequency-domain feature extraction. This makes the proposed method robust to this type of attack. Third, compared with the NSST used by the method in [16], the CT transform has sparse properties and better detail characterization capabilities. Thus, it can filter or perform specific processing to reduce the noise in the cover image to fewer coefficients, allowing for effective noise suppression and thereby improving the ability to resist noise attacks. Fourth, unlike the DTCWT used in [17], the proposed method constructs features by introducing the CT transform, which enables the extraction of more stable principal component information. When subjected to noise, filtering, and JPEG compression, the CT transform can effectively remove high-frequency

Table 15
Comparative experimental results for four different datasets.

Type of attack	BossBase		BOWS-2		COVID		SIPI	
	Average NC	STD	Average NC	STD	Average NC	STD	Average NC	STD
JPEG compression (QF = 5)	0.9899	0.0185	0.9962	0.0038	0.9971	0.0036	0.9963	0.0036
JPEG compression (QF = 15)	0.9973	0.0052	0.9989	0.0012	0.9991	0.0014	0.9990	0.0013
Median filtering (3 × 3)	0.9985	0.0020	0.9992	0.0008	0.9996	0.0008	0.9991	0.0005
Median filtering (11 × 11)	0.9933	0.0069	0.9961	0.0026	0.9985	0.0020	0.9965	0.0030
Wiener filtering (3 × 3)	0.9997	0.0005	0.9998	0.0003	0.9999	0.0002	0.9999	0.0004
Wiener filtering (11 × 11)	0.9979	0.0013	0.9987	0.0009	0.9995	0.0009	0.9994	0.0006
Gaussian low-pass filtering (3 × 3)	0.9999	0.0003	0.9999	0.0002	0.9999	0.0001	0.9999	0.0002
Gaussian low-pass filtering (11 × 11)	0.9998	0.0004	0.9996	0.0003	0.9995	0.0006	0.9996	0.0003
Mean filtering (3 × 3)	0.9985	0.0009	0.9987	0.0008	0.9993	0.0010	0.9982	0.0010
Mean filtering (11 × 11)	0.9946	0.0026	0.9956	0.0021	0.9977	0.0028	0.9961	0.0026
Gaussian noise (0.1)	0.9665	0.0309	0.9846	0.0114	0.9881	0.0116	0.9812	0.0212
Gaussian noise (0.5)	0.9528	0.0244	0.9510	0.0302	0.9609	0.0318	0.9542	0.0245
Salt & pepper noise (0.1)	0.9786	0.0240	0.9912	0.0065	0.9932	0.0070	0.9802	0.0390
Salt & pepper noise (0.5)	0.9537	0.0230	0.9617	0.0267	0.9710	0.0269	0.9622	0.0253
Speckle noise (0.1)	0.9948	0.0041	0.9946	0.0035	0.9962	0.0036	0.9900	0.0017
Speckle noise (0.5)	0.9868	0.0084	0.9845	0.0081	0.9889	0.0088	0.9859	0.0106
Gaussian white noise (0.1,0.05,0)	0.9750	0.0329	0.9936	0.0088	0.9927	0.0095	0.9815	0.0291
Gaussian white noise (0.5,0.25,0)	0.9558	0.0079	0.9714	0.0242	0.9765	0.0252	0.9744	0.0214
Rotation attack (10°)	0.9695	0.0086	0.9703	0.0071	0.9824	0.0120	0.9844	0.0140
Rotation attack (80°)	0.9694	0.0083	0.9716	0.0076	0.9644	0.0304	0.9714	0.0242
Scaling attack (Shrink 0.25x)	0.9990	0.0012	0.9994	0.0010	0.9991	0.0014	0.9991	0.0010
Scaling attack (Magnify 4x)	0.9991	0.0009	0.9990	0.0012	0.9999	0.0002	0.9998	0.0003
Cropping attack (Upper left 1/16)	0.9963	0.0053	0.9988	0.0012	0.9971	0.0032	0.9969	0.0028
Cropping attack (Upper left 1/8)	0.9766	0.0072	0.9859	0.0106	0.9872	0.0186	0.9845	0.0081

Table 16
Experimental results of the proposed method and six similar methods.

Type of attack	Method [16]	Method [17]	Method [18]	Method [19]	Method [22]	Method [25]	Proposed method
JPEG compression (QF = 15)	0.983	0.985	0.995	0.989	0.997	0.996	0.998
Median filtering (3 × 3)	0.996	0.989	0.998	0.980	0.972	1.000	0.999
Wiener filtering (3 × 3)	0.998	0.995	1.000	0.999	0.996	1.000	1.000
Gaussian low-pass filtering (3 × 3)	0.999	0.999	1.000	0.996	0.998	1.000	1.000
Mean filtering (3 × 3)	0.997	0.995	0.995	0.989	0.992	0.979	0.998
Gaussian noise (0.1)	0.987	0.979	0.985	0.981	0.966	0.944	0.991
Salt & pepper noise (0.1)	0.962	0.939	0.964	0.997	0.959	0.954	0.993
Speckle noise (0.1)	0.969	0.966	0.974	0.987	0.971	0.980	0.997
Gaussian white noise (0.1, 0.05, 0)	0.988	0.925	0.984	0.960	0.958	0.940	0.993
Rotation attack (10°)	0.899	0.939	0.890	0.896	0.985	0.985	0.984
Scaling attack (Shrink 0.25x)	0.997	0.995	1.000	0.981	0.992	0.997	1.000
Cropping attack (Upper left 1/16)	0.998	0.996	0.976	0.998	1.000	1.000	0.992
Offset attack (Shift up 2 rows)	0.98	0.969	0.977	0.981	0.973	0.950	0.995

Table 17
Summary of improvement rates from Table 16.

Type of attack	Method [16]	Method [17]	Method [18]	Method [19]	Method [22]	Method [25]	Average
JPEG compression (QF = 15)	0.910%	1.114%	0.910%	1.012%	0.706%	0.706%	0.893%
Median filtering (3 × 3)	0.909%	1.835%	0.706%	2.567%	2.884%	−0.100%	1.467%
Wiener filtering (3 × 3)	0.908%	0.806%	0.000%	0.806%	0.806%	0.000%	0.555%
Gaussian low-pass filtering (3 × 3)	0.100%	0.100%	0.000%	0.402%	0.200%	0.000%	0.134%
Mean filtering (3 × 3)	0.504%	0.302%	0.302%	0.910%	0.605%	2.675%	0.883%
Gaussian noise (0.1)	0.405%	1.226%	0.814%	0.916%	2.588%	4.757%	1.784%
Salt & pepper noise (0.1)	3.115%	5.751%	3.008%	3.762%	3.545%	4.088%	3.878%
Speckle noise (0.1)	2.890%	3.209%	3.746%	4.180%	2.678%	1.735%	3.073%
Gaussian white noise (0.1,0.05,0)	1.120%	7.701%	1.223%	3.438%	4.088%	5.638%	3.868%
Rotation attack (10°)	9.821%	9.333%	10.562%	10.438%	0.306%	0.204%	6.777%
Scaling attack (Shrink 0.25x)	0.705%	0.908%	0.000%	1.833%	0.908%	0.705%	0.843%
Cropping attack (Upper left 1/16)	−0.601%	0.405%	0.303%	0.609%	−0.800%	−0.800%	−0.147%
Offset attack (Shift up 2 rows)	3.323%	2.683%	1.842%	2.577%	3.323%	4.737%	3.081%

signals while retaining the low-frequency signals that represent the cover image, resulting in a more stable extracted feature vector. In summary, our method is robust against most attacks compared to similar methods.

Based on the data in Table 16, the improvement rate of the proposed method compared to the other methods is given in Table 17. It can be seen that, for most attacks, the proposed method outperforms various techniques with an average improvement rate of approximately 2%, indicating that the proposed method is effective.

5.10. Ablation experiment

In this study, a zero-watermarking method that combines CT, CHFMs, and FFT is proposed. The experimental results show that it provides excellent performance. In general, CT is a multi-scale transform that can resist noise and filtering attacks. However, it is difficult to adaptively adjust due to the fixed orientation of its basis functions, resulting in limited adaptive capability against geometric attacks such as rotation. CHFMs utilize the rotational and scaling invariance of

Table 18
Results of ablation experiments.

Type of attack	Our method	Without CT	Without CHFMs	Without FFT
JPEG compression (QF = 15)	0.998	0.973	0.988	0.981
Median filtering (3 × 3)	0.999	0.964	0.998	0.991
Wiener filtering (3 × 3)	1.000	0.968	0.999	0.997
Gaussian low-pass filtering (3 × 3)	1.000	0.989	0.999	0.990
Mean filtering (3 × 3)	0.998	0.950	0.998	0.989
Gaussian noise (0.1)	0.991	0.884	0.968	0.988
Salt & pepper noise (0.1)	0.993	0.798	0.975	0.977
Speckle noise (0.1)	0.997	0.827	0.986	0.967
Gaussian white noise (0.1,0.05,0)	0.993	0.827	0.977	0.974
Rotation attack (10°)	0.984	0.982	0.893	0.964
Scaling attack (Shrink 0.25x)	1.000	0.997	0.902	0.970
Cropping attack (Upper left 1/16)	0.992	0.981	0.881	0.906

moments to resist geometric distortion attacks; however, their global integration property makes them highly sensitive to local distortions, such as compression and noise. FFT-based global spectral analysis enhances robustness to geometric attacks and resists interference in the frequency domain, but it is weak against localized cropping attacks.

To verify how CT, CHFMs, and FFT enhance robustness in our method, we performed ablation experiments. The experimental results are shown in Table 18. It can be seen that these three transforms are complementary in their ability to resist attacks. CT provides resistance to noise and filtering attacks through multi-scale frequency domain features; CHFM provides resistance to geometric attacks, such as rotation and scaling, through geometric invariant features; and FFT enhances resistance to conventional and geometric attacks through frequency domain stability. These three transformations can provide resilience against different types of attacks separately, and their synergistic effect together enhances the overall robustness of our method.

5.11. Complexity comparison

Table 19 summarizes the average running time of the seven methods for processing 100 images under the same experimental conditions. It can be seen that methods [16,17,25] have the shortest running time because they are zero-watermarking methods for a single image; methods [18,19] have an increased running time due to the need to perform operations such as fusion and normalization on multiple images. The method [22] has a relatively long running time due to the need to compute image moments, even though it only processes a single image. The proposed method has the longest running time among the seven methods because it combines operations such as multiple images fusion, CT, CHFMs, and FFT. In practice, the runtime of the proposed method is feasible within 60 s on an ordinary personal computer. Taking this into account, the running time of the proposed method is approximately 31.5 s, which is within the acceptable level.

In the experiments, the sizes of the original cover image and the copyrighted image are assumed to be $N \times N$ and $n \times n$, respectively. The proposed method mainly consists of the following steps: image fusion, CT, CHFMs, FFT, copyrighted image encryption, and zero-watermarking generation. The computational complexities of these steps are $O(4N^2)$, $O\left(\frac{1}{4}N^2\right)$, $O\left(\frac{1}{8}N^3\right)$, $O\left(\frac{1}{64}N^2 \log N\right)$, $O(2n^2)$, and $O(n^2)$, respectively. If some details of the method implementation are ignored, the overall computational complexity of the proposed method can be approximated as $O\left(\frac{1}{8}N^3 + \frac{1}{64}N^2 \log N + \frac{17}{4}N^2 + 3n^2\right)$. Accordingly, the computational complexities in [16–19,22], and [25] are $O\left(\frac{192}{64}N^2\right) + O(2n^2)$, $O\left(\frac{205}{64}N^2\right) + O(2n^2)$, $O\left(\frac{624}{64}N^2\right) + O(2n^2)$, $O\left(\frac{624}{64}N^2\right) + O(2n^2)$, $O\left(\frac{1}{32}N^2 + 2N^2 \log N\right) + O(2n^2)$, and $O\left(\frac{192}{64}N^2\right) + O(2n^2)$, respectively. Similarly, the space complexities of the six steps of the proposed method are $O(4N^2)$, $O(2N^2)$, $O\left(\frac{45}{64}N^2\right)$, $O\left(\frac{1}{16}N^2\right)$, $O(3n^2)$, and $O(4n^2)$, respectively. The overall space complexity of the proposed method can be approximately expressed as $\left(\frac{433}{64}N^2 + 7n^2\right)$. The space complexities in [16–19,22], and [25] are $O\left(\frac{73}{16}N^2\right) + O(5n^2)$, $O\left(\frac{513}{64}N^2\right) + O(5n^2)$, $O\left(\frac{593}{64}N^2\right) + O(5n^2)$, $O\left(\frac{657}{64}N^2\right) + O(6n^2)$, $O\left(\frac{193}{64}N^2\right) + O(5n^2)$, and $O\left(\frac{69}{64}N^2\right) + O(5n^2)$, respectively.

In summary, the computational complexity of the proposed method is approximately $O(N^3)$, and the space complexity is $O(N^2)$. The increase in computational complexity of the proposed method compared to similar methods is primarily due to the introduction of image moments, which enhance resistance against geometric attacks. In terms of space complexity, the proposed method is comparable to its counterparts, indicating that the fusion technique effectively mitigates the problem of increasing storage overhead as the number of images increases.

5.12. Key space and sensitivity analysis

A simple image encryption method based on the Lorenz chaotic system and the Fibonacci Q-matrix is proposed to improve the security of the original binary copyrighted images. Next, the security of the proposed image encryption scheme is analyzed in terms of key space and sensitivity.

5.12.1. Key space

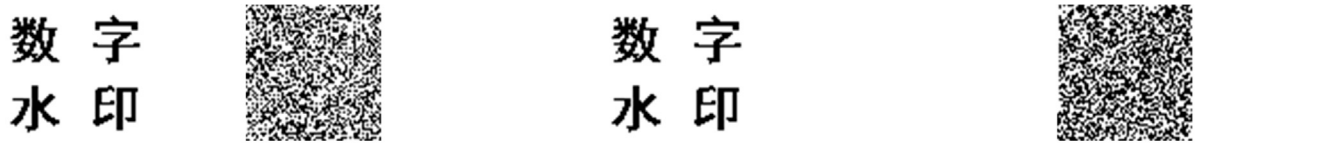
In general, the security of an encryption scheme depends critically on the quantity of its key space. A sufficiently large key space is essential to provide resistance against exhaustive attack. The security of the proposed encryption scheme primarily relies on the initial conditions of the Lorenz chaotic system, as described by Eq. (1). In a 64-bit operating system environment, each parameter is represented as a 64-bit double-precision floating-point number. Consequently, the total key space amounts to $(2^{64})^3 = 2^{192}$. A key space of this magnitude is considered adequate to ensure the cryptographic strength of the encryption scheme against exhaustive attack, thereby enhancing its robustness in practical applications.

5.12.2. Sensitivity analysis

Key sensitivity is regarded as one of the fundamental metrics for evaluating the security of cryptographic schemes. A cryptosystem with high security strength should exhibit significant sensitivity to even minor perturbations in the key. That is, a slight modification in the key should prevent the decryption algorithm from successfully recovering the original plaintext image. The experimental results, depicted in Fig. 12, demonstrate that when the decryption key matches the encryption key precisely, the decrypted image is perfectly consistent with the original. However, when a subtle perturbation is introduced to the decryption key parameter x, i.e., $x_1' = x_1 + 10^{-16}$, the resulting decrypted image becomes severely distorted and entirely unrecognizable to the human eye. This result indicates that the proposed image encryption scheme possesses a high level of key sensitivity, thereby enhancing its resistance against key-related attacks.

Table 19
Comparison of the running times of seven similar methods.

Type of attack	Method [16]	Method [17]	Method [18]	Method [19]	Method [22]	Method [25]	Proposed method
Running time (s)	0.846	1.066	4.326	4.612	5.004	0.907	31.522



(a) Original image (b) Encrypted image (c) Decoded image with the correct key (d) Decrypted image with wrong key

Fig. 12. Experimental results of key sensitivity analysis.

5.13. Discussions

A robust zero-watermarking method is proposed considering the advantages of CT, CHFMs, and FFT. Experimental results show the superior attack resistance of the proposed method against conventional image processing, geometric attacks, and combinatorial attacks. The ablation experimental results show that without CT, the ability to resist noise attacks is weaker; without CHFMs, the ability to resist geometric attacks decreases significantly; and without FFT, the ability to resist noise and cropping attacks decreases slightly. In addition, compared with the methods in [16–19,22,25], our proposed method achieves superior robustness against most attacks. Although these results demonstrate the effectiveness of the proposed method, its limitations remain in the following three aspects.

5.13.1. Ability to resist Gaussian noise

From the experimental results in Tables 5 and 6, it can be concluded that the proposed zero-watermarking scheme exhibits strong robustness against Speckle noise and Salt & pepper noise. However, its performance under Gaussian noise is not satisfactory, indicating a limited resistance to such interference. Consequently, the method's capability to withstand Gaussian noise attacks requires further improvement to enhance its overall robustness.

5.13.2. Low efficiency in calculating CHFMs

From the experimental results in Table 19, it can be concluded that the proposed zero-watermarking method requires approximately 30 s to run on a general-purpose personal computer, indicating that it is not directly applicable to real-time multimedia streaming environments or large datasets. Experiments revealed that the computation of CHFMs constitutes the most time-consuming component in the proposed method, accounting for the majority of the overall execution time. Efficiently computing CHFMs to reduce runtime further is a key issue to be addressed by our method, enabling it to meet real-time requirements.

5.13.3. Scalability of the proposed method

The proposed zero-watermarking generation framework is primarily designed for cover image; therefore, it cannot be directly extended to video watermarking. Evidently, video covers are not only composed of individual frames but also possess inherent relationships between adjacent frames. Applying the proposed technique directly to video scenes often yields unsatisfactory performance. In addition, the zero-watermarking signal generated by the proposed method is stored in a third-party trusted IPR, without considering integration with blockchain technology. The extension of the proposed method to video applications and its integration with blockchain technology would be one of the future research perspectives worthy of in-depth exploration.

6. Conclusion

Aiming to address the limitations of existing zero-watermarking methods, which often exhibit poor performance against specific attacks and can only process a single image, a multi-image robust zero-watermarking method based on CT, CHFMs, and FFT is proposed. First, a high-dimensional chaotic system and a Fibonacci Q-matrix are employed to encrypt a copyrighted image, thereby enhancing the security of the proposed method. Second, multiple images are fused into a single image, and the advantages of the CT, CHFMs, and FFT are combined to construct a feature vector. Numerous experimental results demonstrate that the NC values remain above 0.95 for conventional image processing attacks, geometric attacks, and combined attacks, indicating the proposed method is effective against various types of attacks. Compared to the latest representative methods, it achieves superior performance with an average improvement of approximately 2%. The ablation experiments also confirmed the effectiveness of the combined approach, which utilized CT, CHFMs, and FFT. Although the proposed method can withstand most attacks, its performance still needs improvement. Overall, the limitations of the proposed method are primarily reflected in three aspects. First, the extracted feature vectors are sensitive to noise, resulting in insufficient resilience against attacks such as Gaussian noise. Second, the computational load associated with using CHFMs is high, making it less suitable for real-time applications. Third, the current design is optimized for images and does not directly support videos. To address the limitations above, future work may be focused on the following three perspectives. First, explore the construction of noise-robust feature vectors using advanced feature extraction methods to enhance resistance against noise attacks. Second, improve the computational approach for CHFMs to enhance efficiency, enabling the proposed method to be applied in scenarios with high time-sensitivity requirements. Third, attempt to adapt the proposed method for video by considering its unique spatial and temporal characteristics. Additionally, we plan to integrate blockchain and smart contract technology to create a more comprehensive copyright protection model.

CRediT authorship contribution statement

Xinhui Lu: Writing – original draft, Software, Methodology. **Guangyun Yang:** Visualization, Methodology. **Yu Lu:** Visualization, Methodology. **Xiangguang Xiong:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the Natural Science Foundation of Science and Technology Department of Guizhou Province, China (ZK[2023] 252), the Natural Science Research Project of Guizhou Provincial Department of Education, China ([2023]010), the National Natural Science Foundation of China (U22A2026), and the Qiankehe Platform Talent Foundation of Science and Technology Department of Guizhou Province, China (BQW[2024] 015).

Data availability

Data will be made available on request.

References

[1] Q. Li, B. Ma, X. Wang, C. Wang, S. Gao, Image steganography in color conversion, IEEE Trans. Circuits Syst. II: Express Briefs 71 (1) (2023) 106–110, http://dx.doi.org/10.1109/TCSII.2023.3300330.

[2] T. Yang, H. Wu, B. Yi, G. Feng, X. Zhang, Semantic-preserving linguistic steganography by pivot translation and semantic-aware bins coding, IEEE Trans. Dependable Secur. Comput. 21 (1) (2023) 139–152, http://dx.doi.org/10.1109/TDSC.2023.3247493.

[3] Q. Li, B. Ma, X. Fu, X. Wang, C. Wang, X. Li, Robust image steganography via color conversion, IEEE Trans. Circuits Syst. Video Technol. (2024) http://dx.doi.org/10.1109/TCSVT.2024.3466961.

[4] Y. Chen, A. Malik, H. Wang, B. He, Y. Zhou, H. Wu, Enhancing robustness in video data hiding against recompression with a wide parameter range, J. Inf. Secur. Appl. 83 (2024) 103796, http://dx.doi.org/10.1016/j.jisa.2024.103796.

[5] H. Tao, L. Chongmin, M. Zain, N. Abdalla, Robust image watermarking theories and techniques: A review, J. Appl. Res. Technol. 12 (1) (2014) 122–138, http://dx.doi.org/10.1016/S1665-6423(14)71612-8.

[6] P. Khare, K. Srivastava, A secured and robust medical image watermarking approach for protecting integrity of medical images, Trans. Emerg. Telecommun. Technol. 32 (2) (2021) e3918, http://dx.doi.org/10.1002/ett.3918.

[7] H. Ren, A. Yan, L. Li, Z. Zhang, N. Li, C. Gao, Are you copying my prompt? Protecting the copyright of vision prompt for VPaaS via watermarking, Comput. Stand. Interfaces 94 (2025) 103992, http://dx.doi.org/10.1016/j.csi.2025.103992.

[8] P. Ye, Z. Li, Z. Yang, P. Chen, Z. Zhang, N. Li, J. Zheng, Periodic watermarking for copyright protection of large language models in cloud computing security, Comput. Stand. Interfaces 94 (2025) 103983, http://dx.doi.org/10.1016/j.csi.2025.103983.

[9] Q. Wen, P. Sun, S. Wang, Concept and application of zero watermarking, Acta Automat. Sinica (02) (2003) 214–216, http://dx.doi.org/10.3321/j.issn:0372-2112.2003.02.015.

[10] J. Yang, K. Hu, X. Wang, H. Wang, Q. Liu, Y. Mao, An efficient and robust zero watermarking algorithm, Multimedia Tools Appl. 81 (14) (2022) 20127–20145, http://dx.doi.org/10.1007/s11042-022-12115-8.

[11] C. Chang, C. Chuang, An image intellectual property protection scheme for gray-level images using visual secret sharing strategy, Pattern Recognit. Lett. 23 (8) (2002) 931–941, http://dx.doi.org/10.1016/S0167-8655(02)00023-5.

[12] C. Chang, Y. Lin, Adaptive watermark mechanism for rightful ownership protection, J. Syst. Softw. 81 (7) (2008) 1118–1129, http://dx.doi.org/10.1016/j.jss.2007.07.036.

[13] B. Zou, J. Du, X. Liu, Y. Wang, Distinguishable zero-watermarking scheme with similarity-based retrieval for digital rights Management of Fundus Image, Multimedia Tools Appl. 77 (2018) 28685–28708, http://dx.doi.org/10.1007/s11042-018-5995-4.

[14] C. Wang, D. Qian, D. Ling, Z. Hua, H. Jian, Robust zero-watermarking algorithm via multi-scale feature analysis for medical images, J. Inf. Secur. Appl. 89 (2025) 103937, http://dx.doi.org/10.1016/j.jisa.2024.103937.

[15] N. Ren, Y. Hu, C. Zhu, S. Guo, X. Zhu, Moment invariants based zero watermarking algorithm for trajectory data, J. Inf. Secur. Appl. 86 (2024) 103867, http://dx.doi.org/10.1016/j.jisa.2024.103867.

[16] M. Yang, B. Li, A. Bhatti, Y. Shao, W. Chen, Robust watermarking algorithm for medical images based on non-subsampled Shearlet transform and Schur decomposition, Comput. Mater. Contin. 75 (3) http://dx.doi.org/10.32604/cmc.2023.036904.

[17] T. Huang, J. Xu, Y. Yang, B. Han, Robust zero-watermarking algorithm for medical images using double-tree complex wavelet transform and Hessenberg decomposition, Mathematics 10 (7) (2022) 1154, http://dx.doi.org/10.3390/math10071154.

[18] Y. Lu, H. Lu, Y. Yang, G. Xiong, Robust zero-watermarking algorithm for multi-medical images based on FFST-Schur and Tent mapping, Biomed. Signal Process. Control. 96 (2024) 106557, http://dx.doi.org/10.1016/j.bspc.2024.106557.

[19] B. Wang, W. Wang, P. Zhao, A zero-watermark algorithm for multiple images based on visual cryptography and image fusion, J. Vis. Commun. Image Represent. 87 (2022) 103569, http://dx.doi.org/10.1016/j.jvcir.2022.103569.

[20] X. Wu, J. Li, A. Bhatti, W. Chen, Logistic map and Contourlet-based robust zero watermark for medical images, in: Innovation in Medicine and Healthcare Systems, and Multimedia: Proceedings of KES-InMed-19 and KES-IIMSS-19 Conferences, 2019, pp. 115–123, http://dx.doi.org/10.1007/978-981-13-8566-7_11.

[21] P. Meesala, M. Roy, M. Thounaojam, A robust medical image zero-watermarking algorithm using Collatz and Fresnelet transforms, J. Inf. Secur. Appl. 85 (2024) 103855, http://dx.doi.org/10.1016/j.jisa.2024.103855.

[22] Y. Yang, R. Qi, P. Niu, Y. Wang, Color image zero-watermarking based on fast quaternion generic polar complex exponential transform, Signal Process., Image Commun. 82 (2020) 115747, http://dx.doi.org/10.1016/j.image.2019.115747.

[23] B. Xiao, F. Ma, X. Wang, Image analysis by Bessel–Fourier moments, Pattern Recognit. 43 (8) (2010) 2620–2629, http://dx.doi.org/10.1016/j.patcog.2010.03.013.

[24] G. Gao, G. Jiang, Bessel-Fourier moment-based robust image zero-watermarking, Multimedia Tools Appl. 74 (2015) 841–858, http://dx.doi.org/10.1007/s11042-013-1701-8.

[25] A. Dash, K. Naik, Zero watermarking scheme based on Polar Harmonic Fourier moments, in: International Conference on Computing, Communication and Learning, 2022, pp. 162–171, http://dx.doi.org/10.1007/978-3-031-21750-0_14.

[26] S. Chen, A. Malik, X. Zhang, G. Feng, H. Wu, A fast method for robust video watermarking based on Zernike moments, IEEE Trans. Circuits Syst. Video Technol. 33 (12) (2023) 7342–7353, http://dx.doi.org/10.1109/TCSVT.2023.3281618.

[27] H. Zhu, Y. Yang, Z. Gui, Y. Zhu, Z. Chen, Image analysis by generalized Chebyshev–Fourier and generalized pseudo-Jacobi–Fourier moments, Pattern Recognit. 51 (2016) 1–11, http://dx.doi.org/10.1016/j.patcog.2015.09.018.

[28] C. Gong, J. Liu, M. Gong, B. Li, A. Bhatti, X. Ma, Robust medical zero-watermarking algorithm based on Residual-DenseNet, IET Biom. 11 (6) (2022) 547–556, http://dx.doi.org/10.1049/bme2.12100.

[29] Q. He, Y. He, T. Luo, Y. Song, Shrinkage and redundant feature elimination network-based robust image zero-watermarking, Symmetry 15 (5) (2023) 964, http://dx.doi.org/10.3390/sym15050964.

[30] Y. Liu, C. Wang, M. Lu, J. Yang, J. Gui, S. Zhang, From simple to complex scenes: Learning robust feature representations for accurate human parsing, IEEE Trans. Pattern Anal. Mach. Intell. (2024) http://dx.doi.org/10.1109/TPAMI.2024.3366769.

[31] C. Wang, X. Li, Z. Xia, Q. Li, H. Zhang, J. Li, B. Han, B. Ma, HIWANet: A high imperceptibility watermarking attack network, Eng. Appl. Artif. Intell. 133 (2024) 108039, http://dx.doi.org/10.1016/j.engappai.2024.108039.

[32] Y. Liu, L. Zhang, H. Wu, Z. Wang, X. Zhang, Reducing High-Frequency artifacts for Generative model watermarking via Wavelet transform, IEEE Internet Things J. 11 (10) (2024) 18503–18515, http://dx.doi.org/10.1109/JIOT.2024.3363613.

[33] Y. Liu, H. Wu, X. Zhang, Robust and imperceptible black-box DNN watermarking based on Fourier perturbation analysis and frequency sensitivity clustering, IEEE Trans. Dependable Secur. Comput. (2024) http://dx.doi.org/10.1109/TDSC.2024.3384416.

[34] L. Lin, D. Wu, J. Wang, Y. Chen, X. Zhang, H. Wu, Automatic, robust and blind video watermarking resisting camera recording, IEEE Trans. Circuits Syst. Video Technol. (2024) http://dx.doi.org/10.1109/TCSVT.2024.3448502.

[35] J. Gao, Z. Li, B. Fan, An efficient robust zero watermarking scheme for diffusion tensor-Magnetic resonance imaging high-dimensional data, J. Inf. Secur. Appl. 65 (2022) 103106, http://dx.doi.org/10.1016/j.jisa.2021.103106.

[36] X. Chang, B. Chen, W. Ding, X. Liao, A DNN robust video watermarking method in dual-tree complex wavelet transform domain, J. Inf. Secur. Appl. 85 (2024) 103868, http://dx.doi.org/10.1016/j.jisa.2024.103868.

[37] J. Wang, W. Yu, J. Wang, Y. Zhao, J. Zhang, D. Jiang, A new six-dimensional hyperchaotic system and its secure communication circuit implementation, Int. J. Circuit Theory Appl. 47 (5) (2019) 702–717, http://dx.doi.org/10.1002/cta.2617.

[38] T. Zhou, J. Shen, X. Li, C. Wang, H. Tan, Logarithmic encryption scheme for cyber–physical systems employing Fibonacci Q-matrix, Future Gener. Comput. Syst. 108 (2020) 1307–1313, http://dx.doi.org/10.1016/j.future.2018.04.008.

[39] P. Dong, G. Brankov, P. Galatsanos, Y. Yong, F. Davoine, Digital watermarking robust to geometric distortions, IEEE Trans. Image Process. 14 (12) (2005) 2140–2150, http://dx.doi.org/10.1109/TIP.2005.857263.

[40] H. Song, S. Yu, X. Yang, L. Song, C. Wang, Contourlet-based image adaptive watermarking, Signal Process., Image Commun. 23 (3) (2008) 162–178, http://dx.doi.org/10.1016/j.image.2008.01.005.

[41] Z. Ping, R. Wu, Y. Sheng, Image description with Chebyshev–Fourier moments, J. Opt. Soc. Amer. A 19 (9) (2002) 1748–1754, http://dx.doi.org/10.1364/josaa.19.001748.

[42] Z. Ping, H. Ren, J. Zou, Y. Sheng, W. Bo, Generic orthogonal moments: Jacobi–Fourier moments for invariant image description, Pattern Recognit. 40 (4) (2007) 1245–1254, http://dx.doi.org/10.1016/j.patcog.2006.07.016.

[43] R. Jain, M. Kumar, K. Jain, M. Jain, Digital Image Watermarking using Hybrid DWT-FFT technique with different attacks, in: 2015 International Conference on Communications and Signal Processing, ICCSP, 2015, pp. 0672–0675, http://dx.doi.org/10.1109/ICCSP.2015.7322574.
[44] BossBase image database. https://www.kaggle.com/datasets/lijiyu/bossbase.
[45] BOW-2 image database. https://data.mendeley.com/datasets/kb3ngxfmjw/1.
[46] COVID image database. https://github.com/ieee8023/covid-chestxray-dataset.
[47] SIPI image database. http://sipi.usc.edu/database/.