Scientific Reports

https://doi.org/10.1038/s41598-026-38991-w

Article in Press

A robust zero-watermarking and signcryption scheme for image copyright protection and license verification

Received: 23 December 2025

Accepted: 2 February 2026

Published online: 14 February 2026

Cite this article as: Hung P.T. & Thanh T.M. A robust zero-watermarking and signcryption scheme for image copyright protection and license verification. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-38991-w

Pham Thai Hung & Ta Minh Thanh

A Robust Zero-Watermarking and Signcryption Scheme for Image Copyright Protection and License Verification

Pham Thai Hung[a], Ta Minh Thanh[a,b]

[a]*Institute of Information and Communication Technology, 236 Hoang Quoc Viet, Cau Giay, Ha Noi, VietNam*

[b]*Corresponding author, email: thanhtm@lqdtu.edu.vn*

Abstract— Zero-watermarking (ZW) presents a promising approach for safeguarding image copyright, as it does not alter the original image, a crucial feature for preserving the integrity of medical and high-fidelity visual data. Nevertheless, numerous existing ZW techniques are susceptible to geometric distortions and signal-processing attacks, thereby offering limited protection for ownership and licensing information. This paper proposes a robust and secure zero-watermarking scheme for medical and natural color images that jointly supports ownership authentication and license verification. The method combines entropy- and SIFT-based sub-region selection, DWT-DCT feature extraction, and XOR fusion between robust features and an Arnold-scrambled logo, followed by an ElGamal-style signcryption of the resulting share. Multiple local zero-watermarks are registered in a Certification Authority (CA), enabling global watermark reconstruction without altering the original image. Experimental results show that the normalized correlation (NC) between the recovered and watermark remains above 0.99 under various geometric and non-geometric attacks, confirming the robustness of the scheme. In addition, the signcryption module incurs low computational overhead, with both the encryption and joint decryption–verification processes requiring approximately 8.5 milliseconds. This overhead is small compared with the transform-based processing time and yields a favorable trade-off between enhanced cryptographic protection of ownership/license records and the computational efficiency required for practical medical imaging and large-scale copyright management systems.

Keywords: Zero-watermarking, Scale-Invariant Feature Transform (SIFT), geometric attacks, signcryption, ownership verification, license distribution.

1. Introduction

1.1. Overview

The rapid proliferation of networked multimedia systems and medical imaging platforms has intensified concerns over intellectual property protection and the secure management of digital images. In clinical settings, diagnostic images must maintain visual integrity without perceptible loss, while in broader multimedia applications, high-fidelity content is often shared, archived, and redistributed across diverse platforms. In both scenarios, unauthorized

duplication, alteration, or redistribution of images can result in significant legal, economic, and even clinical ramifications. Digital watermarking has therefore been widely adopted for copyright protection, integrity verification, and anti-counterfeiting, by embedding imperceptible marks into the original image to assert ownership or trace misuse [1, 2].

Conventional image watermarking techniques can be broadly categorized into spatial-domain and transform-domain schemes. Spatial-domain methods directly modify pixel intensities (e.g., LSB substitution or patchwork-based techniques) [3], and are generally simple and low-cost but highly sensitive to common image processing operations such as filtering, compression, noise addition, and so on. Transform-domain methods first map the original image into a frequency or multi-resolution representation, for example via the Discrete Wavelet Transform (DWT) [4, 5], the Discrete Cosine Transform (DCT) [6, 7], or combined multi-transform frameworks. Then, such methods embed watermark bits into selected coefficients [8]. These methods typically offer improved robustness compared with that of spatial-domain schemes, but the embedded watermark can still be degraded or erased by strong compression, geometric transformations, or combined attacks. Furthermore, both spatial- and transform-domain watermarking techniques inherently alter the original image, a modification that is undesirable, and often unacceptable, in safety-critical applications such as medical diagnostics.

In response to these application-driven constraints, recent years have witnessed substantial advances in embedded watermarking and self-recovery watermarking frameworks, particularly for medical imaging and telemedicine applications. A number of studies have demonstrated that carefully designed transform-domain architectures can effectively balance imperceptibility, robustness, payload capacity, and computational efficiency. For example, multiscale schemes integrating the Non-Subsampled Shearlet Transform (NSST) for directional feature extraction, QR decomposition for numerical stability, and adaptive optimization mechanisms such as Particle Swarm Optimization (PSO) have been shown to preserve diagnostic quality while embedding sensitive patient information, without requiring the original image during extraction [9].

Recent works on multimedia security highlight a shift to security-by-design, where watermarking is part of a comprehensive protection framework that includes encryption, authentication, and system-level trust models [10]. Frequency-domain watermarking of biomedical signals and images has been explored to protect electronic patient records in telemedicine, using techniques like Redundant Discrete Wavelet Transform (RDWT) and Schur decomposition for stable, imperceptible embedding of sensitive data [11].

At the algorithmic level, a clear trend has emerged toward hybrid and multi-component watermarking architectures. Several recent studies integrate advanced signal transforms—

such as the Ridgelet Transform, Fractional Discrete Cosine Transform (FDCT), or Mellin Transform—with numerically stable matrix factorizations including QR or Schur decomposition, while employing adaptive Quantization Index Modulation (QIM), bio-inspired optimization algorithms (e.g., Ant Colony Optimization (ACO)), or clustering techniques such as K-means and Self-Organizing Maps (SOM) to reduce perceptual distortion and enhance robustness against noise, filtering, and compression attacks [12–14].

In parallel, state-of-the-art self-recovery watermarking schemes have addressed tamper localization and content restoration through sophisticated block-mapping strategies—such as the Crisscross Block Mapping Strategy (CrCsBMS)—combined with authentication features derived from Gram–Schmidt Orthonormalization (GSO) and multi-stage recovery mechanisms, achieving impressive imperceptibility and payload capacity [15,16].

It is important to emphasize, however, that the aforementioned advances are fundamentally rooted in embedded or self-recovery watermarking paradigms. Despite their strong performance, these methods inherently rely on modifying the original image—either to embed ownership or authentication data, or to enable tamper detection and pixel-level recovery. In safety-critical and high-fidelity imaging scenarios, particularly medical diagnostics, even imperceptible modifications may be undesirable or unacceptable due to strict clinical and regulatory requirements that mandate complete preservation of the original image content.

From the perspective of applications that impose strict content-preservation constraints, an alternative and complementary research direction has emerged in the form of zero-watermarking (ZW), which completely avoids direct embedding. Instead of modifying pixel values, ZW schemes derive a *watermark code* from robust and distinctive image features and register it with a trusted authority. This paradigm is particularly appealing for medical and other high-fidelity images, where any modification of the original content - even if imperceptible - may be unacceptable.

Local feature descriptors such as SIFT [17, 18], SURF [19], or DAISY [20, 21] have been extensively explored in this context due to their invariance to scale, rotation, and moderate geometric distortions. For instance, Fang et al. [17] utilized the SIFT algorithm to extract invariant features from medical images, followed by bandelet transform and DCT to generate the corresponding feature vectors. While this method demonstrates strong robustness against various distortions, it exhibits limited resistance to cropping attacks, incurs high computational costs, and suffers from feature loss under certain geometric transformations such as rotation when pixel values are missing, ultimately leading to incomplete watermark reconstruction. Thanh et al. [22] employed the KAZE feature detector [23] to match feature points between a frame patch and all frames within a video, thereby

identifying the embedding and extraction regions. Similarly, Viet et al. [24] proposed a robust object-based watermarking scheme that integrates SIFT features with a novel data embedding technique based on the DCT domain. Hung et al. [25] proposed a SIFT-based zero-watermarking scheme for robust color image copyright protection by integrating the Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) domains. Salient feature points matching is exploited to estimate rotation, scaling, and translation parameters, enabling geometric normalization of attacked images prior to watermark extraction. Although the method exhibits strong robustness against geometric distortions, it requires storing all extracted feature points in advance, resulting in increased storage overhead. Tsai et al. [26] presented a zero-watermarking approach that combines the Discrete Fourier Transform (DFT) with log-polar mapping to achieve invariance to translation, scaling, and rotation. In the extraction phase, a Support Vector Machine (SVM) optimized by Particle Swarm Optimization (PSO) is employed to estimate the zero-watermark, leading to improved retrieval accuracy at the expense of relatively high computational complexity. Many of these methods demonstrated strong resistance to individual attacks such as rotation, scaling, or moderate compression and have been successfully applied to medical or natural image protection.

However, several important challenges still remain. First, the robustness of many ZW schemes degrades significantly under severe noise contamination and complex combined attacks (e.g., simultaneous rotation, scaling, and translation followed by filtering or compression), leading to incomplete or severely distorted watermark reconstruction. Second, several feature-based methods depend on storing all feature points or patches of the original image to facilitate pre-alignment or recovery before zero-watermark extraction, resulting in significant storage and computational overhead. Third, from a security perspective, most existing ZW approaches protect the logo only by simple permutation-based scrambling (e.g., Arnold transform, chaotic maps, block scrambling) followed by XOR with robust features, and then store the resulting code in plaintext in a certification authority (CA) database. This design leaves the overall scheme to watermark-recovery attacks, ownership forgery, and manipulation of ownership or license records if the CA or the registered codes are compromised.

Furthermore, the majority of ZW schemes are primarily designed for ownership verification and offer limited support for managing usage licenses in multi-party scenarios. In practice, it is often necessary not only to prove the rights of the original copyright owner but also to verify whether an end-user holds a legitimately distributed license (e.g., in medical image sharing, telemedicine services, or commercial content distribution). Ensuring that ownership and license information are cryptographically linked to the zero-watermark

records, while preserving the integrity of the original image, is a critical requirement that has yet to be adequately addressed.

1.2. Our contributions

Based on the above observations, this work investigates a zero-watermarking framework that aims to enhance robustness against strong conventional and geometric attacks while maintaining an acceptable computational cost for practical deployment. In addition, the framework strengthens the cryptographic protection of watermark as well as ownership and license information.

The proposed method leverages entropy- and SIFT-based sub-region selection, DWT–DCT feature extraction, and an ElGamal-like public-key signcryption mechanism to generate and securely store multiple local zero-watermarks for medical and natural color images. To the best of our knowledge, existing studies have not reported an integration of zero-watermarking with a public-key signcryption framework in the manner proposed here for jointly supporting ownership authentication and license verification.

Together with the demonstrated robustness against strong geometric and signal-processing attacks and the low cryptographic overhead, the proposed approach represents a distinctive and practically meaningful contribution to secure image copyright protection.

The main contributions of this paper are summarized as follows:

(1) A sub-region-based zero-watermarking scheme with invariant features.

We introduce a zero-watermarking method that selects the most informative sub-regions using entropy ranking and SIFT keypoints, and then transforms the Y component of each selected sub-region from the DWT-DCT domain. The low-frequency coefficients are binarized to construct robust feature matrices, which are combined with an Arnold-scrambled binary logo by using XOR to generate multiple local zero-watermarks, enabling reliable reconstruction of the global watermark without altering the original image.

(2) An efficient SIFT keypoint selection strategy for RST-attacked image recovery.

We present a method for determining the minimum number of SIFT keypoints essential for accurately recovering images that have undergone RST (Rotation, Scaling, and Translation) transformations. This approach eliminates the need to store the entire set of feature points, thus optimizing storage requirements while maintaining robustness against such geometric attacks. This approach enables effective recovery of images that have undergone both simple and complex attacks commonly encountered in image attacks.

(3) A secure signcryption-based protection of zero-watermark shares and license information.

To address the security limitations of conventional ZW schemes that store the plaintext XOR codes (Zero-Watermark or Ownership Share) in the CA database, we signcrypt the

XOR result $S = MS \oplus W_k$ using an ElGamal-style public-key scheme. MS is Master Share of original image and W_k is the watermark. The resulting pairs (R, C) are bound to the original copyright owner (licensor) and the licensed user (licensee), thereby enhancing confidentiality of the watermark and providing integrity and authenticity for both ownership and license records, while mitigating watermark-recovery and ownership/license-manipulation attacks.

(4) Comprehensive robustness and efficiency evaluation against a state-of-the-art scheme.

We conduct extensive experiments on medical and standard color images under Gaussian noise, median filtering, JPEG compression, rotation, scaling, translation, and cropping. The proposed scheme consistently achieves higher NC values than that of method of S.A. Nawaz et al. [27], with NC ≈ 1.0 for many geometric and signal-processing attacks. In addition, execution-time measurements show that the cryptographic operations incur only a few milliseconds (ms) per image and that zero-watermark generation and extraction remain computationally feasible, confirming the suitability of the proposed scheme for practical medical image protection and digital copyright management.

1.3. Roadmap

The remainder of this paper is organized as follows: Section 2 provides the theoretical background. Section 3 presents our proposed zero-watermarking scheme. Section 4 reports the experimental results and corresponding analyses. Finally, Section 5 concludes the paper.

2. Preliminaries

2.1. Scale-invariant feature transform (SIFT)

The scale-invariant feature transform (SIFT) [28] detects distinctive local keypoints that are robust to scale and rotation changes. First, a scale-space is constructed by applying Gaussian smoothing at multiple scales, and extrema are located in the difference-of-Gaussian (DoG) images. Unstable points are removed by checking contrast and edge response criteria. For each remaining keypoint, a dominant orientation is assigned based on local gradient distributions. Finally, a 128-dimensional descriptor is formed by aggregating gradient histograms in a neighborhood around the keypoint, providing robustness to illumination and moderate geometric distortions.

2.2. Entropy analysis

In this work, entropy is employed to quantify the information content of the SIFT-based salient regions extracted from the original image. Intuitively, entropy measures the degree of randomness and uncertainty in the intensity distribution of a region; higher entropy implies richer structural details and, hence, a more informative feature area. Instead of computing entropy over the whole image, we evaluate it locally on patches centered at SIFT keypoints,

so that only highly informative regions are selected for subsequent processing. Mathematically, the Shannon entropy of a discrete random variable (X) associated with the gray-level (or feature) distribution is defined as [29]:

$$H(X) = -\sum_x p(x) \log p(x) \quad (1)$$

where $p(x)$ denotes the probability of occurrence of value (x) within the considered SIFT region.

2.3. DWT, DCT, and Arnold transform

2.3.1. Discrete Wavelet Transform – DWT

The discrete wavelet transform (DWT) [30] is used to perform a multi-resolution decomposition of the image, separating its low- and high-frequency components. At each level, the image is filtered by low-pass and high-pass analysis filters along rows and columns, followed by downsampling, yielding the LL, LH, HL, and HH subbands. The LL subband concentrates most of the signal energy and is typically exploited for robust feature extraction or watermark representation. Thanks to its joint spatial–frequency localization and multi-scale nature, DWT offers improved robustness to common signal-processing operations compared with purely spatial-domain approaches.

2.3.2. Discrete Cosine Transform – DCT

The discrete cosine transform (DCT) [31] is employed as an orthogonal transform to compact most of the image energy into a small number of low-frequency coefficients. In typical image-processing applications, the image is partitioned into non-overlapping blocks (e.g., 8×8), and each block is transformed from the spatial domain to the cosine domain. The DC coefficients and a few low-frequency AC coefficients capture the main structural content, whereas higher-frequency coefficients mainly describe fine details and rapid intensity changes. For robustness-oriented watermarking and feature extraction, selected low- and mid-frequency coefficients are usually exploited, since they are less sensitive to common signal-processing operations and moderate compression (e.g., JPEG).

2.3.3. Arnold transform

In watermarking schemes, to enhance the algorithm's security, the watermark information is often scrambled using the Arnold transform [32, 33] before being embedded into the original image. This technique repositions a pixel from its original coordinates (x,y) to new coordinates (s,t), as represented by the following equation.

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \ (mod\ N) \quad (2)$$

where N × N is the size of the watermark image and *mod*(.) is the modulus operation. To recover the watermark, the inverse Arnold transform is applied as follows:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} \ (mod\ N) \tag{3}$$

The scrambling key, determined by the number of iterations, plays a decisive role in the security level of the Arnold transform.

2.4. The ElGamal-style encryption-authentication scheme

The ElGamal–style encryption–authentication scheme, originally introduced by [34], is a signcryption-type variant of the standard ElGamal public-key algorithm defined over finite fields, which is capable of simultaneously providing confidentiality and authentication (of both origin and integrity) for the encrypted message. This signcryption scheme combines origin authentication and ciphertext integrity directly in the encryption process, unlike traditional methods that require an additional digital signature for authentication.

Furthermore, by employing this framework, the public key of the buyer can be utilized to encrypt the digital product, ensuring that only the buyer, with their corresponding private key, can decrypt the product and verify their ownership rights. This approach streamlines the process of ownership verification, enhancing both security and efficiency.

The ElGamal–style encryption–authentication scheme presented in this section includes: the Key Generation algorithm (Algorithm 1), the Encryption algorithm (Algorithm 2) and the Decryption – Authentication algorithm (Algorithm 3).

2.4.1. Signcryption

Signcryption is a class of public-key primitives that simultaneously provides encryption and digital signature functionality within a single logical operation. Instead of performing "signature-then-encryption" or "encryption-then-signature" as two separate steps, a signcryption scheme combines both processes into one integrated algorithm. The concept was first formalized by Zheng in 1997, with the central design principle that the overall cost of a signcryption operation should be strictly lower than that of the naive combination of a public-key encryption scheme and a digital signature scheme [35]. This principle is often summarized as Eq. 4:

$$\mathrm{Cost}(\text{Signature \& Encryption}) \ll \mathrm{Cost}(\text{Signature}) + \mathrm{Cost}(\text{Encryption}) \tag{4}$$

This inequality can be interpreted in several complementary ways. First, a signcryption scheme should be more efficient in terms of computational complexity than the straightforward composition of encryption and signature. Second, it should produce a more compact ciphertext–tag output, i.e., the signcrypted text must be shorter than the concatenation of a conventional public-key ciphertext and a separate digital signature. Third, beyond efficiency and compactness, a well-designed signcryption scheme may also offer

stronger or strictly better-integrated security guarantees (e.g., combined confidentiality, authenticity, and non-repudiation) than the naive two-step approach [35]. The practical importance of signcryption has been increasingly recognized in modern secure communication systems.

2.4.2. The parameter and Key generation algorithm

In this work, DSA-style parameter and key generation are used to instantiate the ElGamal-type encryption–authentication core. The Digital Signature Algorithm (DSA) is a standardized public-key signature scheme based on the discrete logarithm problem in a prime field. It was originally introduced by the U.S. National Institute of Standards and Technology (NIST) as part of the Digital Signature Standard (DSS) in FIPS PUB 186 [36].

Algorithm 1: Generate parameters and keys

input: l_p, l_q

output: p, q, g, y, x

1: Choose a pair of prime numbers p, q with:

$\text{len}(p) = l_p$, $\text{len}(q) = l_q$ and $q \mid (p-1)$,

where (l_p, l_q) is typically chosen from $(1024, 160)$, $(2048, 224)$, $(2048, 256)$ or $(3072, 256)$, respectively.

2: Choose a value of α in the range $(1, p)$, compute g according to the formula:

$g = \alpha^{\frac{p-1}{q}} \ mod \ p$, satisfying $g \neq 1$.

3: Choose a secret key x in the range $(1, q)$.

4: Calculate the public key y according to the formula:

$y = g^x \ mod \ p$.

Notes:

- len(): The function that calculates the length (in bits) of an integer.
- y: The public key.
- x: The secret (private) key.
- p, q, g: The system parameters.

Assume x_s is the secret key of the sender/encryptor and x_r is the secret key of the receiver/decryptor, then the corresponding public key of the sender is:

$$y_s = g^{x_s} \ mod \ p$$

And that of the receiver is:

$$y_r = g^{x_r} \ mod \ p$$

2.4.3. The Encryption algorithm

Algorithm 2: Encryption

input: p, g, x_s, y_r, P

output: (R, C)

1: Compute the value S_e according to the formula:

$$S_e = (y_r)^{x_s} \ mod \ p$$

2: Compute the value R by:

$$R = HASH(P)$$

3: Compute the sender's encryption key K_e by:

$$K_e = HASH(R \parallel S_e)$$

4: Encrypt the plaintext P according to the formula:

$$C = P * g^{K_e} \ mod \ p$$

5: Send ciphertext $(R, C)$to the receiver.

Notes:

- y_r: The public key of the receiver.
- x_s: The secret (private) key of the sender.
- P: The plaintext.
- (R, C): The ciphertext corresponding to P.
- $HASH()$: The cryptographic hash function, e.g. $SHA1/SHA256$.
- Operator ‖ is the operation to concatenate two bit-strings.

2.4.4. The Decryption - Authentication algorithm

Algorithm 3: Decryption - Authentication

input: $p, g, x_r, y_s, (R, C)$

output: M

1: Compute the value S_d according to the formula:

$$S_d = (y_s)^{x_r} \ mod \ p$$

2: Compute the receiver's decryption key K_d by:

$$K_d = HASH(R \parallel S_d)$$

3: Decrypt the received ciphertext C according to the formula:

$$M = C * g^{-K_d} \ mod \ p$$

4: Compute the value V by:

$$V = HASH(M)$$

5: Checks: if $V = R$ then the origin and integrity of the post–decrypted message M is confirmed. Otherwise, if $V \neq R$, the validity of the received message will be denied.

Notes:

- y_s: The public key of sender.
- x_r: The secret key of the receiver.

- M: The post–decrypted message.

3. Our proposed method

In this section, we present our proposed zero-watermarking and signcryption–based framework for robust image copyright protection and license distribution. The key idea is to exploit the inherent robust features of prominent local sub-regions in the original image and to bind them, via signcryption, to a cryptographically protected ownership code.

First, SIFT is applied to detect stable local keypoints, and for each keypoint a surrounding patch (sub-region) is extracted. The use of SIFT ensures invariance to scale and rotation transformations, allowing the same salient regions to be reliably detected even under geometric distortions. The information content of these patches is then quantified using the local Shannon entropy defined in Section 2.2. Entropy is employed as a complementary criterion to characterize the richness and complexity of local textures, so that sub-regions with insufficient structural information are excluded. As a result, only the sub-regions with the highest entropy values are retained as the most informative and robust feature areas.

Based on the invariant features of the selected sub-regions, we construct a set of sub-region zero-watermarks. Each zero-watermark is then processed using the proposed ElGamal-style signcryption mechanism (Section 2.4) to ensure the confidentiality and authenticity of the embedded copyright data.

During the verification phase, the same SIFT- and entropy-based selection procedure is performed on the queried image. The corresponding sub-region zero-watermarks are reconstructed, and the global watermark is reassembled from the set of verified sub-region watermarks. This global watermark, together with the signcryption-based authentication, enables reliable image copyright verification and supports secure management and distribution of usage licenses.

3.1. Sub-region zero-watermark generation

Let the original binary watermark image be denoted as **W** with size $N \times N$, and a color image $\mathbf{R}_i$ with size $2N \times 2N$, as the sub-region image of original image **I**. Fig.1 illustrates the sub-region zero-watermark generation process. The steps of this procedure are presented in *Algorithm* 4:

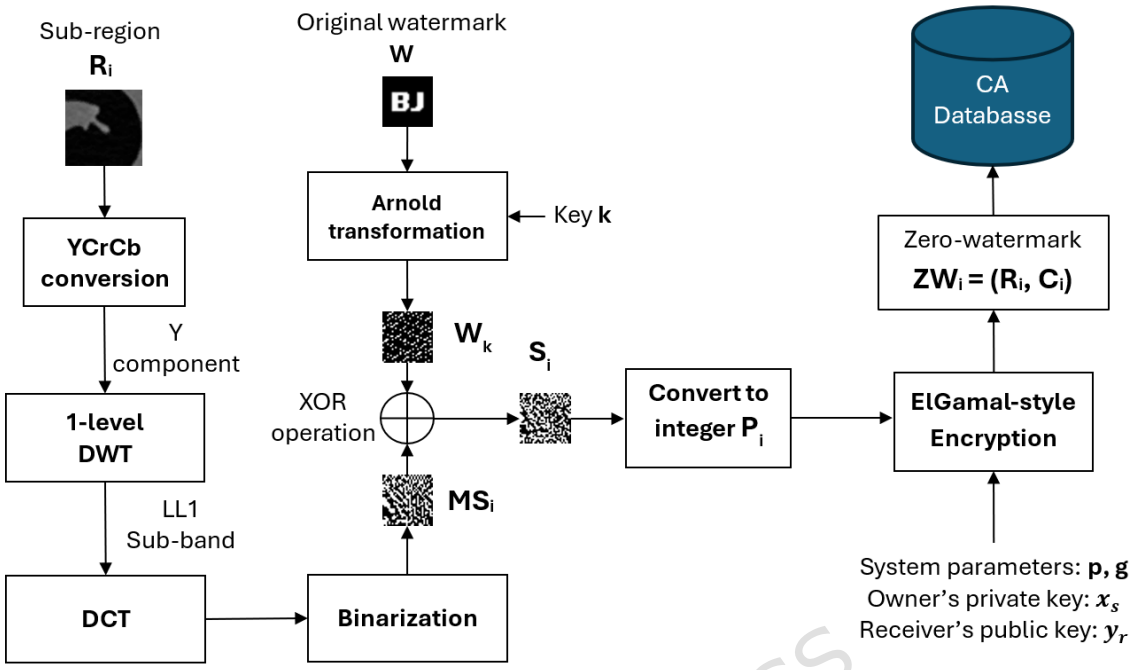

Fig. 1. Flow chart of the Sub-region zero-watermark generation process

---

Algorithm 4: Sub-region Zero-Watermark Generation

Input: Sub-region image $\mathbf{R_i}$, watermark image $\boldsymbol{W}$, system parameters **p, g,** owner's private key $\boldsymbol{x_s}$ and receiver's public key $\boldsymbol{y_r}$.

Output: Sub-region zero-watermark images $\boldsymbol{ZW_i}$.

1: Convert $\mathbf{R_i}$ to YCbCr domain to obtain Y-component;

2: Apply 1-level DWT to Y-component to obtain LL1 sub-band;

3: Perform DCT to LL1 sub-band to obtain DCT_LL1 coefficient matrix;

4: Binarize the DCT_LL1 to obtain the master share $\boldsymbol{MS_i}$;

5: Scramble the watermark **W** using the Arnold transform with a secret key k to obtain the encrypted watermark $\mathbf{W_k}$;

6: $\mathbf{S_i} \leftarrow \mathbf{MS_i} \oplus \mathbf{W_k}$.

7: $\mathbf{P_i} \leftarrow$ ConvertBinaryMatrixToInterger($\mathbf{S_i}$)

8: Encrypt the $\mathbf{P_i}$ using **Algorithm 2** with system parameters **p, g,** owner's private key $\boldsymbol{x_s}$ and receiver's public key $\boldsymbol{y_r}$ to obtain the sub-region zero-watermark $\mathbf{ZW_i} = \mathbf{(R_i, C_i)}$

9: $\mathbf{ZW_i}$ is stored in the CA database for subsequent processing when proving the original owner, and can be delivered to the license purchaser in the case of copyright license distribution.

In the proposed framework, the Certification Authority (CA) is assumed to operate under a semi-trusted trust model. The CA is responsible for storing and managing registered zero-

watermarks for copyright verification in dispute scenarios. However, the CA is not trusted with respect to the confidentiality or integrity of ownership and license information, as all zero-watermark records are protected using the proposed public-key signcryption scheme. Consequently, even if the CA is compromised, the underlying watermark content and ownership claims cannot be forged or disclosed without the corresponding cryptographic keys. This trust model and its security implications are discussed in more detail in Section 4.5.

3.2. Sub-region watermark extraction

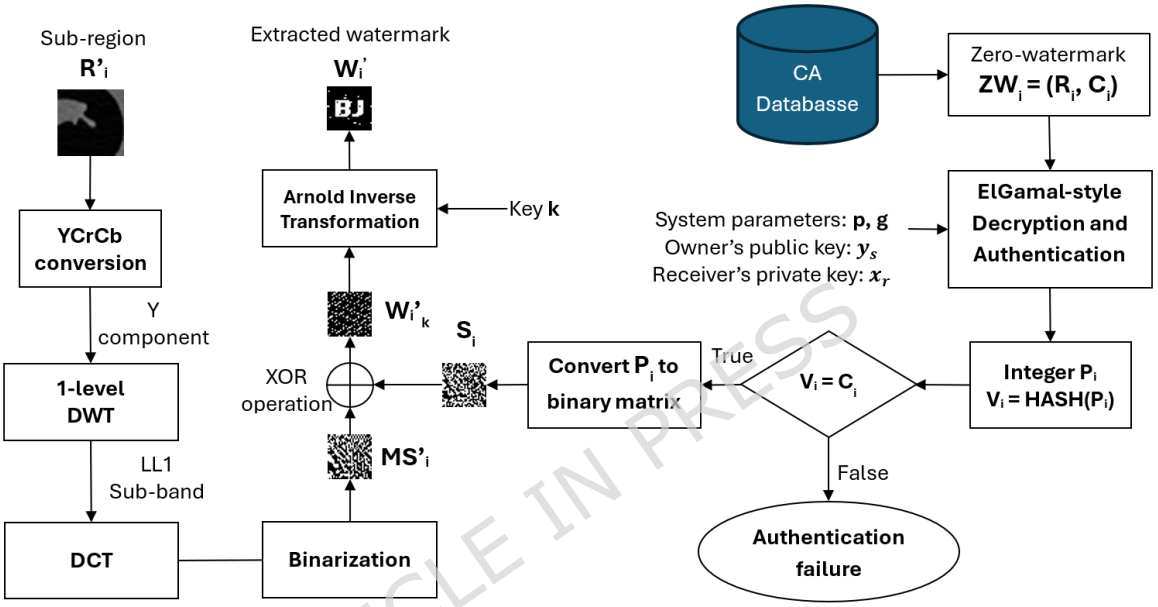

Fig. 2. Flow chart of the Sub-region watermark extraction process

Fig. 2 illustrates the sub-region zero-watermark extraction process on a sub-region color image **R'i** with size 2N×2N. The steps of this procedure are presented in Algorithm 5:

Algorithm 5: Sub-region Watermark Extraction

Input: Sub-region image **R'i**, sub-region zero-watermark images ZW_i, system parameters **p, g,** owner's public key y_s and receiver's private key x_r

Output: Sub-region watermark image W_i'

1: Convert **R'i** to YCbCr domain to obtain Y-component;

2: Apply 1-level DWT to Y-component to obtain LL1 sub-band;

3: Perform DCT to LL1 sub-band to obtain DCT_LL1 coefficient matrix;

4: Binarize the DCT_LL1 to obtain the master share MS'_i;

5: Decrypt the $ZW_i = (R_i, C_i)$ using **Algorithm 3** with system parameters **p, g,** owner's public key y_s and receiver's private key x_r to obtain P_i;

6: V_i = HASH(P_i);

7: **If** ($V_i \neq C_i$) **then**

8: Decryption–authentication failure;
9: **return;**
10: $\mathbf{S}_i \leftarrow$ ConvertIntergerToBinaryMatrix($\mathbf{P}_i$)
11: $\mathbf{W_i'}_k \leftarrow \mathbf{MS'_i} \oplus \mathbf{S}_i$.
12: Apply the Arnold transform with key k on $\mathbf{W_i'}_k$ to obtain $\mathbf{W_i'}$

3.3. Image correction

Geometric attacks such as rotation, scaling, and translation remain significant challenges for watermarking algorithms in general and for zero-watermarking schemes in particular. To improve the robustness of zero-watermarking, we consider a new pipeline to automatically estimate the parameters of geometric attacks for correcting image before the watermark extraction process is carried out. Unlike many existing zero-watermarking approaches that require access to the original image for attack compensation, our method avoids increased memory usage and computational overhead. Instead, image recovery is achieved using only a small set of prominent SIFT keypoints and their descriptors that were previously extracted and stored in the CA database. In addition, our approach eliminates the need to recover images subjected to translation attacks. The main steps are shown as follows:

Step 1: Firstly, we perform Step 1-Step 2 in Section 3.4 to extract the top **P** strongest SIFT keypoints and their descriptors from the original image **I**, called **(K_i,D_i),** i=1,…,**P**. They are stored in CA database for further matching process. In our experiments, the parameter **P** was set to 50.

Step 2: Read the keypoints and their descriptors saved from original image **I** and match them and those of the attacked image **I'** to find out the matched points.

Step 3: We estimate the Homography matrix **H** between the two sets of matched points using RANSAC algorithm:

$$\boldsymbol{H} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} ($$

Step 4: Calculate the RS (Rotation, Scaling) parameters:

$$\boldsymbol{A} = \boldsymbol{H} * \boldsymbol{H} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{bmatrix} \quad (5)$$

- Rotation angle: $\boldsymbol{\alpha} = atan2(h_{10}, h_{00}) * 180/\pi$ (6)

- Scale fator: $\boldsymbol{\beta} = \sqrt{\frac{a_{00}+a_{01}+a_{10}+a_{11}}{2.0}}$ (7)

Step 5: Finally, the attacked image **I'** is recovered using the estimated RS parameters.

3.4. Global Zero-watermark generation process

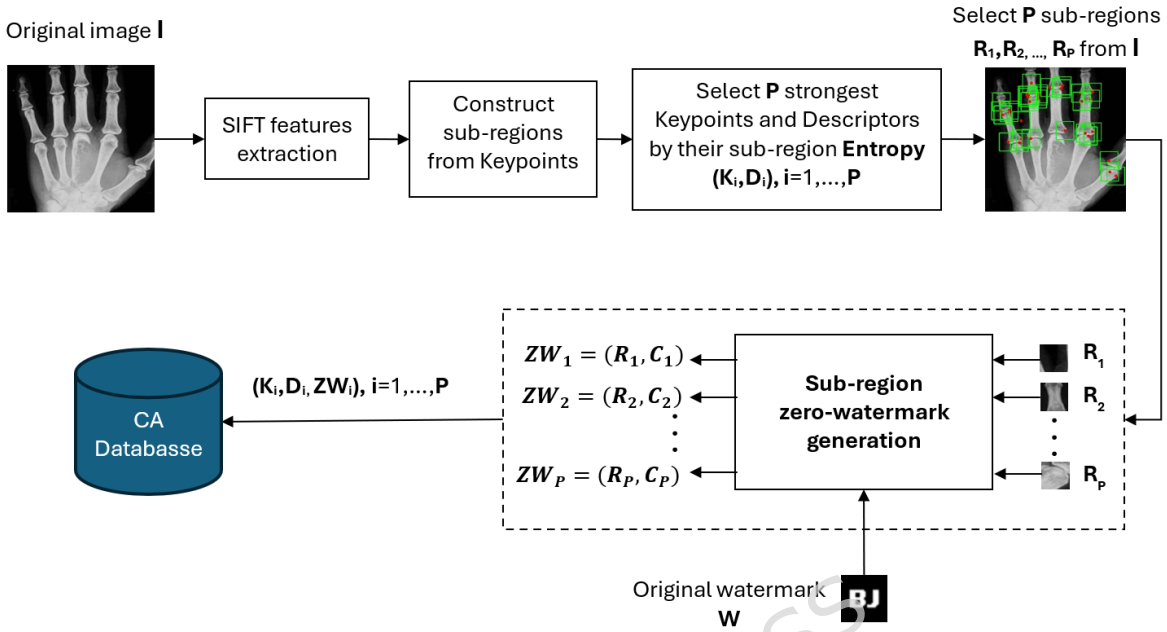

Fig. 3. Flow chart of the Global Zero-watermark generation process

A binary image **W** with size N×N, as the original watermark, and a color image **I** with size M×M, as the original image, are chosen to prove our zero-watermark generation. In our experiments, the **W** and **I** are 32×32 and 512×512 pixels, respectively. **Fig. 3** shows the zero-watermark generation process. The steps of this procedure are presented in Algorithm 6:

Algorithm 6: Zero-watermark Generation

Input: Original color image **I**, original watermark image $\boldsymbol{W}$, system parameters **p, g,** owner's private key $\boldsymbol{x_s}$ and receiver's public key $\boldsymbol{y_r}$

Output: $\boldsymbol{P}$ tuples of sub-region zero-watermarks $\mathbf{ZW_i}$ and the keypoints $\mathbf{K_i}$ with their descriptors $\mathbf{D_i}$, $\mathbf{(K_i, D_i, ZW_i)}$

1: Extract all SIFT keypoints and their descriptors of the original image **I**;
2: **foreach** extracted keypoint $\mathbf{K_i}$
3: Construct sub-region $\mathbf{R_i}$ from the image **I** with size of 2N×2N pixels, $\mathbf{K_i}$ is the center point of $\mathbf{R_i}$;
4: Calculate the entropy value $\mathbf{E_i}$ of the sub-region $\mathbf{R_i}$ using Eq. (1) in Section 2.2;
5: **end foreach**
6: Sort the list $\{\boldsymbol{R_i}\}$ in descending order of their entropy values
7: Select the top $\boldsymbol{P}$ sub-regions together with their corresponding keypoints and descriptors, $\mathbf{(K_i, D_i, R_i)}$, i=1,…,**P**;
8: **for** i=1 to **P do**
9: Calculate $\mathbf{ZW_i}$ of the sub-region $\mathbf{R_i}$ using **Algorithm 4**;

10: **end for**

11: Store ***P*** tuples of **(K_i, D_i, ZW_i)** in CA database

3.5. Zero-watermark verification process

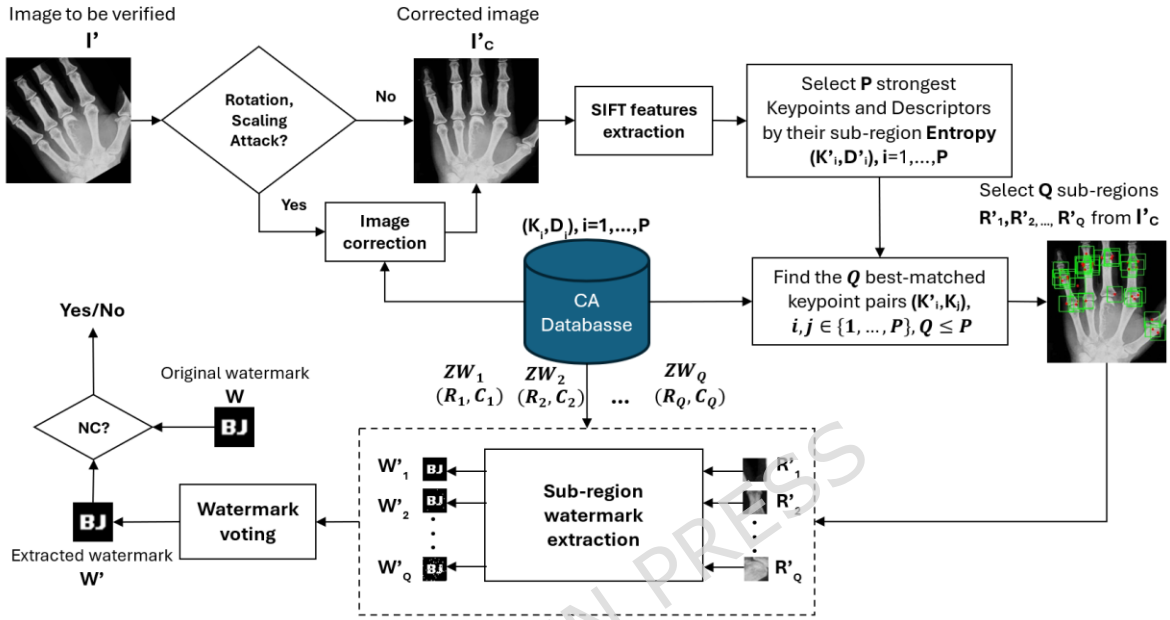

Fig. 4. Flow chart of the Global Zero-watermark verification process

Our zero-watermark verification process is used to verify the watermark of potentially attacked image **I'**. The pipeline of our zero-watermark verification has been illustrated in **Fig. 4** and the main steps are presented in Algorithm 7. The parameter **P** and threshold **T** are established based on our experimental results.

In order to reconstruct the global watermark **W'**, a voting-based strategy is applied to the extracted sub-region watermarks, according to the following approach:

$$\boldsymbol{W'}(p,q) = \begin{cases} \mathbf{1}, if \ \sum \boldsymbol{W'}_i(p,q) \geq \frac{\boldsymbol{Q}}{2} \\ \mathbf{0}, otherwise \end{cases} \quad (8)$$

where $\boldsymbol{W'}_i$ are extracted sub-region watermarks and $\boldsymbol{Q}$ is the overall count of the best-matched keypoints.

Algorithm 7: Zero-watermark Verification

Input: Potential attacked image **I'**, original watermark image ***W***, ***P*** tuples of **(K_i, D_i, ZW_i)** in CA database and Euclidean similarity threshold **T,** system parameters **p, g,** owner's public key $\boldsymbol{y_s}$ and receiver's private key $\boldsymbol{x_r}$

Output: Copyright conclusion: Yes/No

1: **if I'** has undergone geometric attacks (e.g., rotation or scaling) **then**

2: Perform Image Correction method in Section 3.3 to obtain the corrected image **I'C**;

3: **end if**
4: Extract all SIFT keypoints and their descriptors of the corrected image $\mathbf{I'_C}$ and arrange in descending order based on their sub-regions' entropy values;
5: Select the top $\boldsymbol{P}$ strongest extracted keypoints and their descriptors, $\mathbf{(K'_i, D'_{i'})}$, i=1,…,**P**;
6: $\boldsymbol{Q}$=0; index=0;
7: **for** i=1 to **P do**
8: max_similarity = 0;
9: **for** j=1 to **P do**
10: similarity = Euclidean_ similarity ($\mathbf{D'_i}$, $\mathbf{D_j}$);
11: **if** similarity > max_similarity **then**
12: max_similarity = similarity;
13: index = j;
14: **end if**
15: **end for**
16: **if** max_similarity >=**T then**
17: $\boldsymbol{Q} = \boldsymbol{Q}+1$;
18: Construct sub-region $\mathbf{R'_i}$ from the image $\mathbf{I'_C}$ with size of 2N×2N pixels, $\mathbf{K'_i}$ is the center point of $\mathbf{R'_i}$;
19: Extract $\mathbf{W'_i}$ using **Algorithm 5** with $\mathbf{R'_i}$ and sub-region zero-watermark $\boldsymbol{ZW}_{\mathbf{index}}$;
20: **end if**
21: **end for**
22: Calculate the global watermark **W'** from $\boldsymbol{Q}$ sub-region watermarks $\mathbf{W'_i}$ using Eq. (8);
23: Calculate the **NC** value between **W** and **W'** to judge the copyright of **I'**;

Code availability: The core implementation of the proposed zero-watermarking framework is publicly available at: https://github.com/hungpt-mta/zero-watermarking-core.

4. Evaluation and experimental results

Our experiments were carried out on a Windows-based platform with the following configuration: Windows 11 Home Single Language (version 24H2), 32 GB RAM, and an Intel(R) Core(TM) Ultra 7 155H CPU @ 1.40 GHz. All implementations were written in Python and executed in a Jupyter Notebook environment. In particular, we used Python 3.12.7 together with several auxiliary libraries for data processing and result analysis.

For the evaluation, we selected 12 medical images at random from the public repository TCIA (https://nbia.cancerimagingarchive.net/nbia-search/) and 3 color images from the widely used USC-SIPI standard image dataset (http://sipi.usc.edu/database/). All test images have a spatial resolution of 512×512 pixels. The watermark logo is a binary image of size

32×32, containing information derived from the trial set. These images were used to assess both the effectiveness and the robustness of our proposed watermarking algorithm. The test images and the watermark are illustrated in Fig. 5.

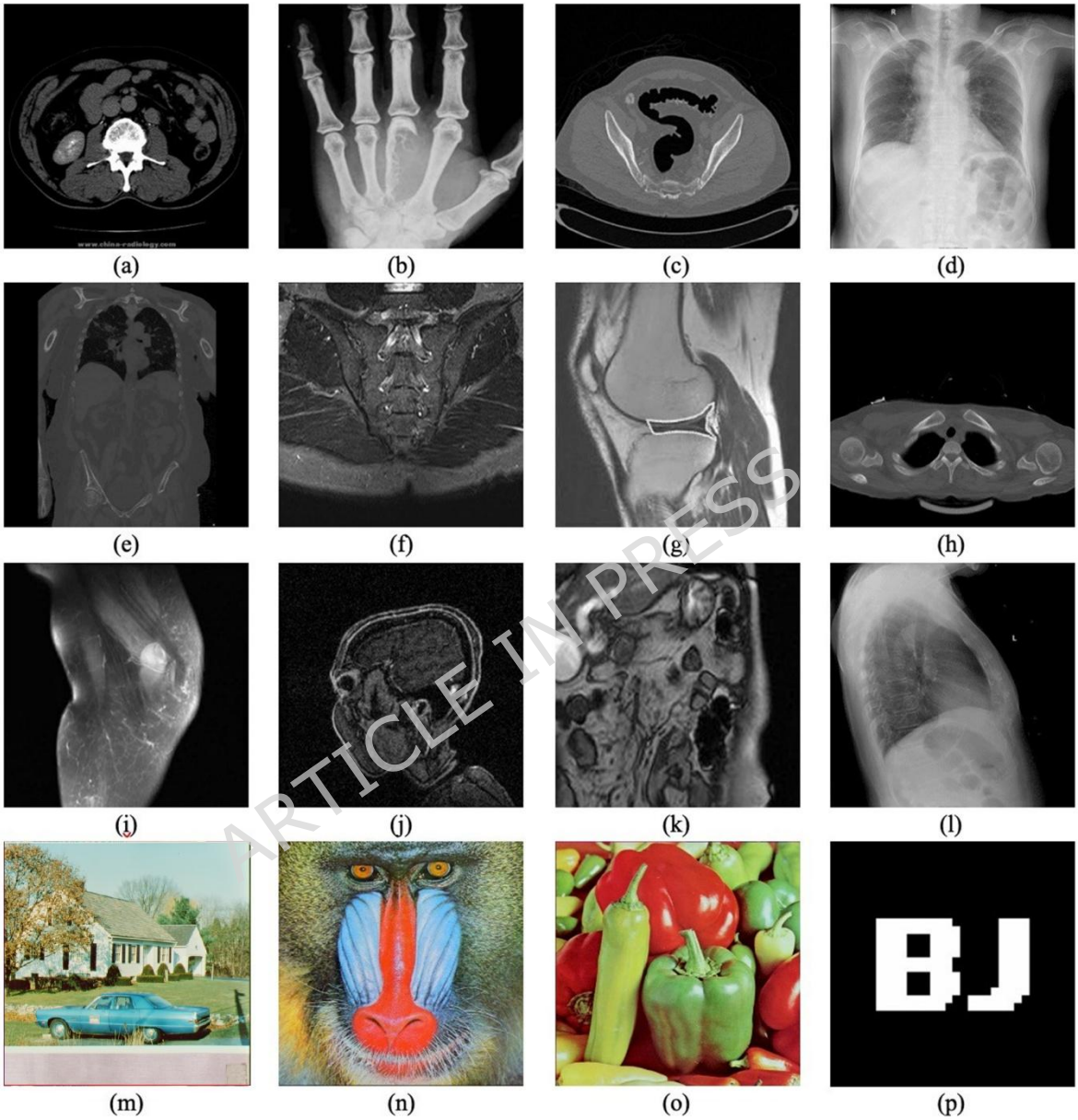

Fig. 5. 12 medical images from TCIA (a-l), 3 color images from USC-SIPI (m-o) and the binary watermark logo (p)

4.1. Evaluation metrics

In this work, the robustness of the proposed scheme is evaluated in terms of the Normalized Correlation (NC) and the Peak Signal-to-Noise Ratio (PSNR). The NC index quantifies the similarity between the original watermark and the extracted one. The larger NC values indicate a stronger correlation and, consequently, higher robustness of the

watermarking method, as defined in Eq. (9). The PSNR metric, in contrast, characterizes the amount of distortion introduced into the watermarked image. A low PSNR implies that the watermarked image deviates significantly from the original, whereas a high PSNR value corresponds to better preserved visual quality. The PSNR is computed according to Eq. (10).

$$NC = \frac{\sum_i \sum_j W_{(i,j)} W'_{(i,j)}}{\sum_i \sum_j W^2_{(i,j)}} \tag{9}$$

where $W_{(i,j)}$, $W'_{(i,j)}$, denotes the original watermark and the extracted watermark, respectively.

$$PSNR = 10log\left[\frac{MN \max_{i,j}(I_{(i,j)})^2}{\sum_i \sum_j (I_{(i,j)} - I_{(i,j)}')^2}\right] \tag{10}$$

where $I_{(i,j)}$ is the gray-level value of the pixel at position (i, j) in the original and watermarked images, M and N are the height and width (in pixels) of the medical image, respectively.

To further examine the effectiveness and practical relevance of the proposed algorithm, we perform a comparative study against the method of S.A. Nawaz et al. [27]. In this experiment, both schemes are subjected to the same set of image processing and geometric attacks, and the corresponding zero-watermarks are extracted. The resulting NC values are then compared to assess the relative robustness and detection accuracy of the two approaches.

From a research standpoint, such a comparison is crucial because the NC provides a quantitative measure of how reliably the watermark can be preserved under adverse conditions. If our method consistently yields higher NC values than the reference scheme, this indicates not only stronger resistance to a wide range of attack scenarios but also greater reliability in real-world settings, where images are routinely exposed to compression, noise, and geometric transformations. Additionally, these findings indicate that the proposed algorithm strikes an optimal balance between robustness and computational efficiency, fulfilling a key design objective in modern digital watermarking systems.

4.2. Analysis and Optimization of Parameters P and T

In this experiment, we first extract the complete set of SIFT keypoints from each test image. The original images are then subjected to a range of typical image-processing and geometric attacks, such as Gaussian noise addition, median filtering, and JPEG compression. For every attacked version, the image is reconstructed using the procedure described in Section 3.3, and the corresponding zero-watermark is obtained. The similarity between the extracted zero-watermark and the original one is quantified by the Normalized Correlation (NC), which serves as the main performance indicator.

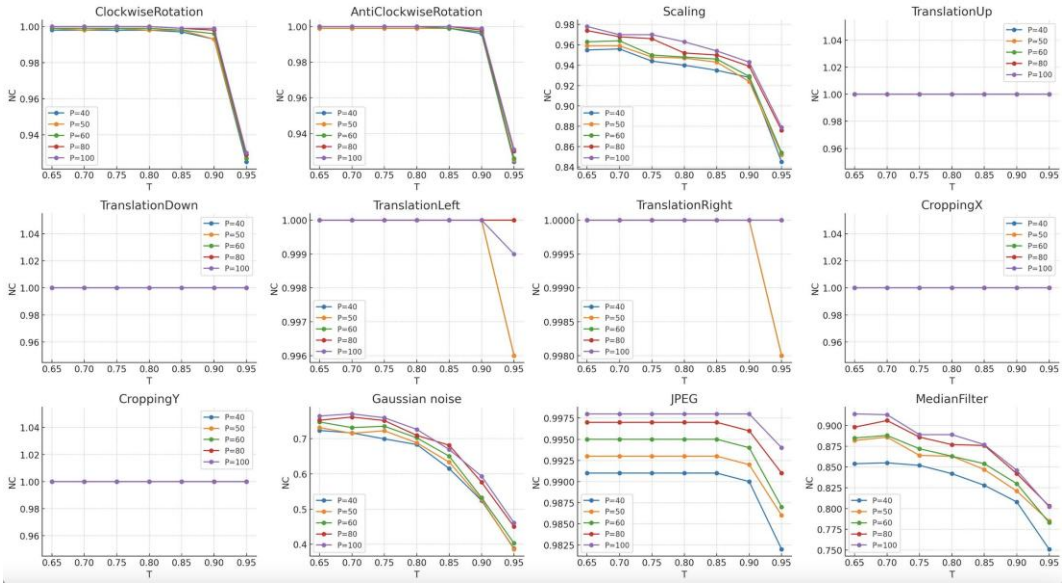

Fig. 6. Impact of the number of SIFT keypoints (**P**) and the matching threshold (**T**) on the robustness of the proposed scheme.

To analyze the impact of the parameters, we compare the SIFT feature vectors of the original and attacked images for different numbers of preserved keypoints $P \in \{40,50,60,80,100\}$ and matching thresholds $T \in \{0.65,0.70,0.75,0.80,0.85,0.90,0.95\}$, as summarized in Fig. 6.

The results show that increasing P tends to slightly improve the NC values, implying better robustness in watermark reconstruction. However, the improvement from $P = 40$ to $P = 100$ is not substantial, indicating that beyond a certain point, preserving more keypoints yields only marginal benefits. On the other hand, raising the threshold T leads to a clear decrease in NC because a stricter matching condition eliminates many candidate correspondences, especially when the images are strongly distorted.

Taking into account the balance between robustness, computational complexity, and storage overhead, we adopted $P = 50$ and $T = 0.65$ as the recommended parameter setting. With these values, the algorithm automatically selects sub-regions whose SIFT descriptors remain the most stable under the considered attacks (see Fig. 7). These sub-regions form a reliable basis for zero-watermark extraction across diverse image processing and geometric distortions. By concentrating on the most resilient local structures, the proposed method maintains high watermark integrity without embedding any modifications into the original image, thereby achieving both imperceptibility and strong resistance to adversarial operations.

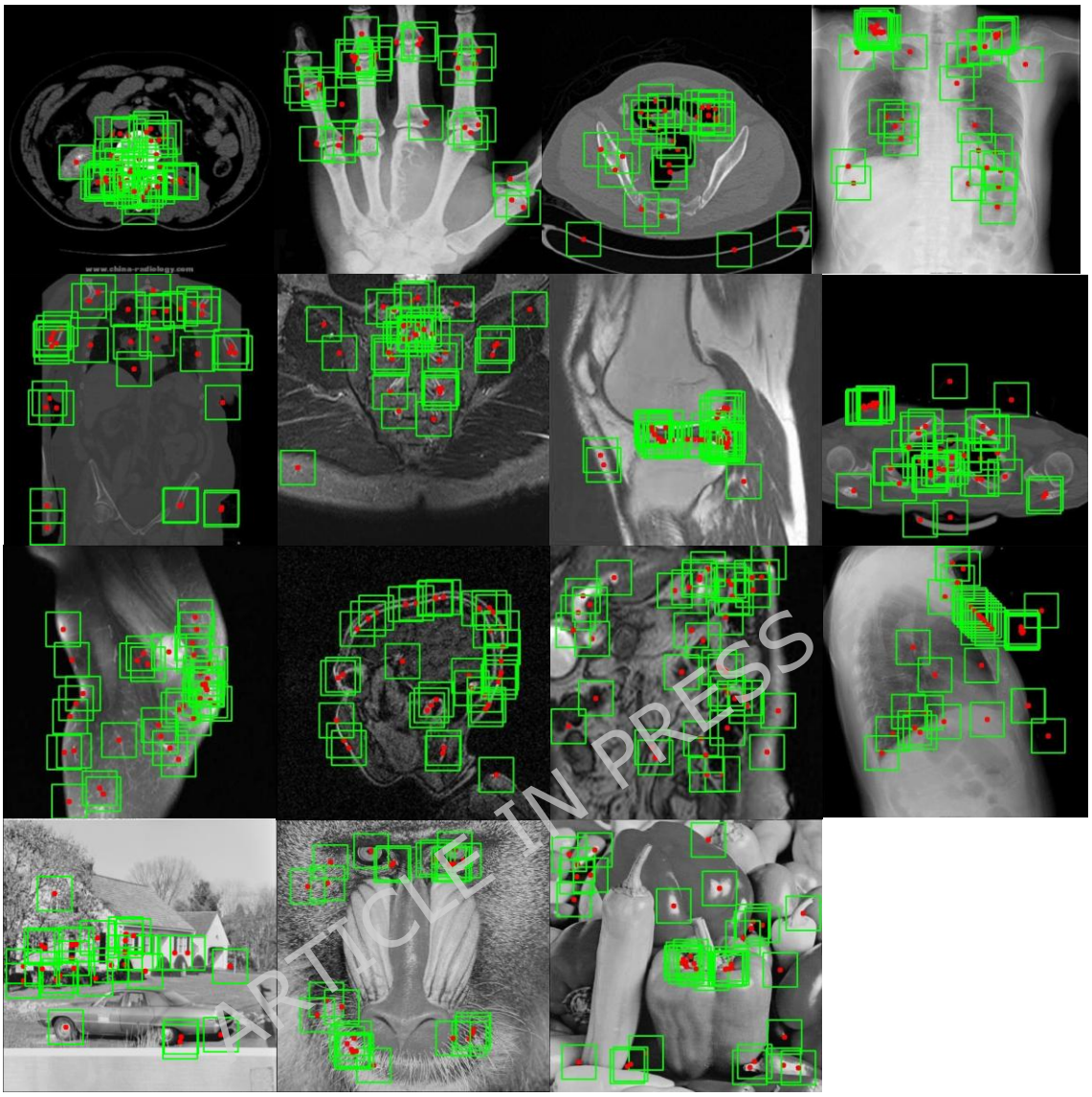

Fig. 7. Sub-regions extracted from the experimental image datasets based on SIFT feature keypoints and entropy magnitude.

Table 1 summarizes the NC values obtained by the proposed scheme and by the method of S.A. Nawaz [27] across various attack conditions. For a fair and reliable comparison, both approaches were evaluated on exactly the same set of medical images and their associated watermarked counterparts, and the results for our scheme correspond to the parameter configuration (P = 50) and (T = 0.65), so that any differences in performance are not influenced by variations in image content.

Table 1: Comparison of robustness under various attacks

Attacks	Intensity	PSNR (dB)	Nawaz et al.	Proposed
Gaussian Noise	1%	22.00	0.78	**0.82**
	2%	19.10	0.72	**0.76**
	5%	15.37	0.54	**0.68**
JPEG Compression	15%	31.80	0.95	**1**
	25%	33.86	1	**1**
	30%	34.53	1	**1**
	40%	35.14	0.95	**0.99**
	50%	36.52	1	**0.99**
Median Filter (Average)	[3x3]	29.16	0.75	**0.99**
	[5x5]	25.78	0.71	**0.90**
	[7x7]	23.61	0.63	**0.77**
Rotation clockwise (°)	10°	16.68	0.80	**1**
	20°	15.33	0.80	**1**
	60°	13.92	0.87	**1**
	70°	13.67	0.91	**1**
	80°	13.50	0.95	**1**
Rotation Anticlockwise (°)	15°	15.86	0.80	**1**
	30°	14.89	0.80	**1**
	50°	14.27	0.85	**1**
	60°	13.92	0.85	**1**
	80°	13.50	0.90	**1**
Scaling	× 0.4	---	0.77	**0.87**
	× 0.6	---	0.73	**0.98**
	× 0.9	---	1	**1**
	× 1.2	---	---	**1**
	× 1.4	---	---	**1**
Left Translation (%)	10%	13.91	1	**1**
	15%	13.06	1	**1**
	20%	12.65	0.90	**1**
	30%	12.24	0.86	**1**
	35%	12.20	0.87	**1**
Right Translation (%)	10%	13.98	0.96	**1**
	15%	13.10	0.83	**1**
	20%	12.73	0.86	**1**
	30%	12.20	0.73	**1**
	35%	12.11	0.78	**1**
Up Translation (%)	10%	13.79	0.95	**1**
	15%	12.94	1	**1**
	20%	12.27	0.96	**1**
	25%	11.78	0.92	**1**
	30%	11.42	0.87	**1**

Down Translation (%)	7%	14.75	0.85	**1**
	10%	13.97	0.81	**1**
	15%	13.07		**1**
	20%	12.42		**1**
	25%	11.97		**1**
Cropping (X-axis)	3%	---	0.90	**1**
	10%	---	0.86	**1**
	15%	---	0.90	**1**
	20%	---	0.82	**1**
	25%	---	0.72	**1**
Cropping (Y-axis)	3%	---	0.80	**1**
	10%	---	0.85	**1**
	15%	---	0.90	**1**
	20%	---	0.90	**1**
	25%	---	0.71	**1**

4.3. Robustness analysis against Conventional attacks

4.3.1. Gaussian noise attack

As summarized in Table 1, the robustness of the proposed method is evaluated under additive Gaussian noise with three noise levels of 1%, 2%, and 5%, and its performance is compared with that of the scheme by S.A. Nawaz et al. [27]. The reported NC values consistently reveal a clear advantage of the proposed method at all noise intensities.

At a noise level of 1%, the proposed method achieves a Normalized Correlation (NC) of 0.82, whereas the method of S.A. Nawaz attains only 0.78. When the noise level is increased to 2%, the NC obtained by the proposed method remains high at 0.76, in contrast to 0.72 for the reference method. Under the most severe condition with 5% Gaussian noise, both algorithms show a decrease in NC; however, our method still maintains a relatively strong NC of 0.68, which is markedly higher than the 0.54 achieved by S.A. Nawaz's scheme.

These results demonstrate that the proposed zero-watermarking method exhibits substantially improved tolerance to random Gaussian noise. This enhanced robustness underscores its suitability for real-world applications, particularly in medical image protection and copyright enforcement, where noise contamination is commonly introduced during image acquisition, transmission, and storage.

4.3.2. JPEG compression attack

As reported in Table 1, the robustness of the proposed zero-watermarking scheme is further investigated under JPEG compression with quality settings corresponding to compression ratios of 15%, 25%, 30%, 40%, and 50%. The obtained NC values are compared with those method of S.A. Nawaz et al. [27].

At a compression ratio of 15%, the baseline scheme of S.A. Nawaz achieves an NC of 0.95, whereas the proposed method attains an NC of 1.0, indicating perfect watermark reconstruction. When the compression ratio is increased to 25% and 30%, both methods still yield optimal performance with NC = 1.0. At 40% compression, however, the baseline approach begins to degrade, with the NC dropping to 0.95 and noticeable distortion appearing in the extracted watermark. In contrast, the proposed method maintains an NC of 0.999, and the recovered watermark remains clearly discernible. Under the most severe compression level of 50%, our method produces an NC of 0.995, while the scheme of S.A. Nawaz attains 1.0; despite this slight deviation, the watermark extracted by the proposed method remains highly consistent and visually intact.

Collectively, these results demonstrate that the proposed zero-watermarking technique provides stable and reliable performance even under JPEG compression attacks, making it highly suitable for copyright protection in situations where lossy storage or transmission is inevitable.

4.3.3. Median filter attack

The robustness of the proposed watermarking scheme is further evaluated under median filtering attacks with kernel sizes of [3 x 3], [5 x 5], and [7 x 7]. The performance is compared with the method of S.A. Nawaz et al. [27], considering both the NC values and the visual fidelity of the extracted watermark.

With a [3×3] kernel, the baseline method attains an NC of 0.75, and the recovered watermark exhibits pronounced structural distortions. In contrast, the proposed method achieves an NC of 1.0, indicating perfect preservation of the watermark. When the kernel size is increased to [5×5], the NC value of the baseline method decreases to 0.71, whereas the proposed method still attains a substantially higher NC of 0.95, with the watermark remaining clearly recognizable despite slight noise artifacts. Under the strongest filtering condition, [7×7], the baseline approach further degrades to an NC of 0.63, leading to severe distortion in the extracted watermark. Meanwhile, the proposed method maintains an NC of 0.87 and preserves the essential structure of the watermark even under strong smoothing, demonstrating markedly superior robustness to median filtering.

4.4. Robustness analysis against Geometric attacks

4.4.1. Clockwise and anticlockwise rotation attack

Table 1 reports the NC values obtained when the watermarked images are subjected to rotation attacks with different angles. The results show that the proposed method consistently produces higher NC values than those of method of S.A. Nawaz et al. [27].

Under both clockwise and counterclockwise rotations, a clear performance gap is observed between the two schemes. For counterclockwise rotations from 15° to 80°, the

method of Nawaz et al. yields NC values in the range of 0.80–0.90, and the extracted watermark becomes increasingly distorted and barely recognizable as the angle grows. In contrast, the proposed method maintains NC = 1.0 for all tested angles, indicating perfect preservation of the watermark. A similar behavior is observed for clockwise rotations between 10° and 80°: the baseline approach degrades progressively, with NC values dropping to between 0.80 and 0.95 and noticeable visual artifacts in the recovered watermark, particularly at larger angles. Meanwhile, the proposed method again exhibits superior robustness, retaining NC = 1.0 (or 0.999 at 70°) and enabling almost lossless reconstruction of the watermark. These results confirm that the proposed method offers markedly improved resistance to rotational attacks, while ensuring stable and visually faithful watermark recovery, thereby establishing a clear advantage over the method of Nawaz et al. [27].

4.4.2. Scaling attack

Table 1 presents the results obtained under scaling attacks. These experiments confirmed the superior robustness of the proposed method compared with the method of S.A. Nawaz et al. [27]. For reduced scale factors of ×0.4 and ×0.6, the method of Nawaz et al. yields NC values of 0.77 and 0.73, respectively, and the corresponding extracted watermarks exhibit pronounced structural distortion. In contrast, the proposed method attains higher NC values of 0.87 and 0.98, with the watermark remaining clearly recognizable, particularly at ×0.6.

At a near-original scale of ×0.9, both approaches achieve NC = 1.0, indicating successful and lossless watermark recovery. However, for enlarged scales of ×1.2 and ×1.4, no results are reported for the baseline method, whereas the proposed scheme consistently maintains NC = 1.0 and continues to extract the watermark without visible degradation. Overall, these findings indicate that the proposed algorithm is not only more robust to downscaling but also reliably preserves watermark integrity under upscaling, thereby demonstrating higher robustness and adaptability than the method of Nawaz et al.

4.4.3. Translation left and right attack

Table 1 reports the results obtained under horizontal translation attacks in both right and left directions and further illustrates the robustness of the proposed watermarking algorithm compared with the method of S.A. Nawaz et al. [27].

For rightward translations between 10% and 35% of the image width, the proposed method consistently achieves a normalized correlation (NC) of 1.0, indicating perfect preservation of the watermark and a recovered pattern that is visually indistinguishable from the original. By contrast, the method of Nawaz et al. degrades as the translation ratio increases: the NC decreases from 0.96 at 10% to 0.73 at 30%, and the extracted watermarks become progressively distorted and unclear.

A similar trend is observed for leftward translations. The proposed method again maintains NC = 1.0 for all tested shifts from 10% to 35%, with the extracted watermark remaining clear and fully recognizable. In comparison, the baseline method exhibits reduced robustness, with NC values dropping to as low as 0.86 at a 30% shift and visible distortions appearing in the recovered watermark.

Taken together, these results demonstrate that the proposed scheme provides excellent resistance to geometric translation attacks in both directions, ensuring accurate watermark recovery even under substantial shifts, whereas the method of S.A. Nawaz shows noticeable vulnerability, with lower NC values and degraded visual quality of the extracted watermark.

4.4.4. Translation up and down attack

Translation attacks are implemented by vertically shifting the image pixels upward or downward by a given ratio, which may adversely affect the embedded watermark. Table 1 compare the performance of the proposed method with that of S.A. Nawaz et al. [27] under such conditions.

Across all tested translation ratios, for both upward and downward shifts, the proposed method consistently achieves perfect watermark recovery with NC = 1.0, indicating that the watermark is fully preserved. In contrast, the method of S.A. Nawaz et al. exhibits lower and more variable NC values, with noticeable degradation in the extracted watermark, particularly at larger shift ratios. These results demonstrate that the proposed method offers markedly improved robustness to geometric translation, maintaining watermark integrity even under severe vertical displacements.

4.4.5. Clipping X and Y-direction attack

Clipping (cropping) attacks are modeled by removing portions of the image along the horizontal (X) or vertical (Y) direction by a specified ratio, which can significantly impair the detectability of the embedded watermark. Table 1 compares the performance of the proposed method with that of the method by S.A. Nawaz et al. [27] under such attacks.

For all tested cropping ratios in both the X and Y directions, the proposed method consistently achieves perfect watermark recovery with NC = 1.0, indicating excellent robustness. In contrast, the method of S.A. Nawaz et al. exhibits decreasing NC values as the cropping ratio increases, leading to progressively degraded extracted watermarks. These results demonstrate the superior resistance of the proposed method to cropping attacks and underscore its reliability in practical scenarios involving image resizing or partial removal, such as medical image processing and online content sharing.

4.5. Security Analysis

4.5.1. Vulnerability to watermark recovery attacks

To enhance the security level of the algorithm, most zero-watermarking schemes employ scrambling techniques such as the Arnold transform, Logistic Map, Block Scrambling, etc., to distort the watermark information W into W_k before combining it (typically using the XOR operation) with a robust feature MS extracted from the original image, in order to generate the zero-watermark ZW which is then directly stored in the CA database. Since the embedding and extraction algorithms in copyright protection are usually public, if an attacker can gain access to the CA database and obtain ZW, they can compute the scrambled watermark W_k as follows:

- If the attacker can access the original image (or has a sufficiently good copy of it), they can extract the exact robust feature MS from this image based on the published algorithm.
- Then, they compute $W_k = ZW \oplus MS$.

A weakness of the above scrambling techniques is that their key space is relatively small and often periodic, so an attacker can exploit this by performing brute-force key search or statistical attacks, etc., to recover the original watermark W from W_k.

In the proposed scheme, after computing $S = MS \oplus W_k$, we do not store S (which plays the role of ZW in traditional schemes) directly in the CA database. Instead, we encrypt S using the signcryption scheme described in Section 2.4.3 with the system parameters (p, q, g), the owner's private key (x_s), and the public key (y_r) (of the owner in the case of ownership verification, or of the buyer in the case of copyright distribution). The encryption produces ZW = (R, C), and only then is ZW stored in the CA database. Therefore, in order to recover the watermark W, an attacker would first have to break the ElGamal-style signcryption algorithm in Section 2.4.3, which is practically infeasible. Furthermore, instead of storing a single global zero-watermark value as in conventional schemes, our algorithm generates and stores multiple local zero-watermarks for different sub-regions and uses them to synthesize and reconstruct the global watermark W. As a result, the probability that an attacker can successfully recover W is significantly reduced.

4.5.2. Ownership and License Manipulation Attacks

For conventional zero-watermarking schemes as described above, if an attacker gains control over the CA database, they can modify the information of the original copyright owner (the owner) or the licensed user (licensee) by altering the mapping records between the OwnerID or BuyerID and the zero-watermark ZW, since ZW does not directly contain any information about the owner or the licensee.

In the proposed scheme, ZW is generated by signcrypting directly with the owner's private key and the owner's public key (in the case of ownership verification) or the licensee's public key (in the case of user license verification). Therefore, if an attacker only changes the mapping between OwnerID or BuyerID and the zero-watermark ZW, the decryption–authentication procedure for copyright verification cannot be successfully carried out because the correct original private–public key pair cannot be used.

In the case where an attacker both modifies the OwnerID or BuyerID information and then creates a new ZW using the published signcryption algorithm with a forged private–public key pair of a fake owner or buyer to replace the original ZW in the CA database, the probability that such a forged ZW can still pass the decryption–authentication procedure is also very low, because the attacker cannot guess the system parameters such as *(p, q, g)*.

4.6. Execution Time Analysis

The computational efficiency of the proposed scheme is evaluated in terms of execution time for four main operations: encryption, zero-watermark generation, decryption and authentication, and watermark extraction. Table 2 reports the corresponding runtimes (in milliseconds) for each test image.

Table 2: Execution time (in ms) of the proposed scheme for each test image

Test images	Encryption	Generate Zero-Watermark	Decryption and Authentication	Extract Watermark
img1.png	7.1	404.3	8.4	1091.4
img2.png	6.9	392	8.6	1097.1
img3.png	7.3	396.3	8.7	1099.6
img4.png	8.9	421.3	8.7	958
img5.png	9.5	453.1	8.8	1005.9
img6.png	8.5	473	8.6	1036
img7.png	8.3	423	8.8	1162.6
img8.png	8.4	400.4	8.7	1093.1
img9.png	9.1	424.4	8.4	1027.7
img10.png	7.3	458.3	9	1250.1
img11.png	11.4	622.8	8.8	1073.9
img12.png	9.5	436.8	9	1024.6
house.png	7.8	421.5	8.4	952.7
mandril.png	7.5	442.1	7.6	939.8
peppers.png	8.2	431.2	8.3	1030.1

For the encryption stage, the execution time ranges from 6.9 ms to 11.4 ms, with an average of approximately 8.4 ms across all images. The decryption and authentication step exhibits similarly low cost, with runtimes between 7.6 ms and 9.0 ms and an average of about 8.6 ms. These results indicate that the cryptographic operations introduce only a minor overhead and are suitable for real-time or near real-time use in practical systems.

By contrast, the zero-watermark generation and extraction stages naturally require more processing time, as they involve transform-domain feature extraction and sub-region processing. The generation of the sub-region zero-watermarks takes on average around 439 ms per image (ranging from 392.0 ms to 622.8 ms), while the watermark extraction step requires approximately 1.06 s on average, with values spanning from 939.8 ms to 1250.1 ms. Despite being the most time-consuming components, these runtimes remain acceptable for typical copyright protection workflows, where watermark registration and verification are usually performed off-line or at moderate request rates.

Overall, the measurements in Table 2 show that the proposed method achieves a reasonable trade-off between robustness, security, and computational cost. The cryptographic layer is lightweight, and the transform-based zero-watermark generation and extraction are efficient enough to support practical deployment in medical image protection and related applications.

5. Conclusion

This paper has proposed a secure and robust zero-watermarking scheme for medical and natural color images that supports both ownership authentication and license verification without modifying the original image. The method integrates entropy- and SIFT-based sub-region selection, DWT–DCT feature extraction, and a secure encoding stage in which an Arnold-scrambled logo is XOR-combined with robust features and then protected by an ElGamal-style signcryption mechanism. Multiple local zero-watermarks are registered at the certification authority, allowing reliable reconstruction of the global watermark while preserving the original content.

Extensive experiments on 12 medical images and 3 standard color images, using NC and PSNR under a wide range of conventional and geometric attacks, show that the proposed scheme consistently outperforms the method of S.A. Nawaz et al. [27], often achieving NC ≈ 1.0 and visually lossless watermark recovery. From a security perspective, signcrypting $S = MS \oplus W_k$ with system parameters (p, q, g), the owner's private key, and the owner's or licensee's public key overcomes key weaknesses of traditional XOR-only zero-watermarking, providing confidentiality, integrity, and authenticity of the watermark records and strong resistance to watermark-recovery and ownership/license-manipulation attacks. Moreover, the measured execution times confirm that the cryptographic overhead is small

and that zero-watermark generation and extraction are computationally feasible for practical copyright protection and medical imaging applications.

Despite these advantages, several limitations of the proposed scheme remain. Although strong robustness is achieved against a wide range of geometric attacks, the resistance to severe noise-based attacks is comparatively less pronounced and could be further improved. In addition, the copyright verification stage currently requires approximately one second per image, which may limit scalability in large-scale or time-sensitive deployment scenarios.

Future work will therefore focus on enhancing robustness against strong noise perturbations by incorporating more noise-resilient feature representations and adaptive feature fusion strategies. Another promising research direction is the optimization of the verification pipeline to reduce extraction time, for example through parallel processing or lightweight feature selection, thereby improving efficiency for large-scale copyright management systems and real-time medical imaging applications.

Funding Declaration

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 04/2025/TN.

Data availability

The medical images used in this study were randomly selected from the publicly available The Cancer Imaging Archive (TCIA) repository and can be accessed at: https://nbia.cancerimagingarchive.net/nbia-search/. In addition, three standard color images were obtained from the widely used USC-SIPI image database, available at: http://sipi.usc.edu/database/. All datasets analysed during the current study are publicly available from the above repositories and were used in accordance with their respective terms of use.

No additional permission or informed consent is required to publish or reuse the images used in this study. The medical images were obtained from The Cancer Imaging Archive (TCIA), which provides publicly available datasets that have been fully de-identified in accordance with applicable ethical and legal standards, including HIPAA. Therefore, the images do not contain any personally identifiable information. The color images were taken from the USC-SIPI standard image database, which is a publicly available benchmark dataset widely used for research and reproducibility purposes. Consequently, all images used in this study can be safely employed for reproducibility without ethical or consent-related restrictions.

Code availability

The core implementation of the proposed zero-watermarking framework is publicly available at: https://github.com/hungpt-mta/zero-watermarking-core.

The provided code is sufficient to run the benchmarking procedures described in this paper.

The experimental datasets are not redistributed within the code repository, but are publicly available from The Cancer Imaging Archive (TCIA) and the USC-SIPI image database, and can be obtained directly from their respective sources.

References

[1] P. T. Nha, T. M. Thanh, and N. T. Phong, "Consideration of a robust watermarking algorithm for color images using improved QR decomposition," Soft Computing, vol. 26, no. 11, pp. 5069–5093, Jun. 2022, doi: 10.1007/s00500-022-06975-3.

[2] M. Iwakiri and T. M. Thanh, "Incomplete cryptography method using invariant Huffman code length for digital rights management," in Proc. IEEE Int. Conf. Advanced Information Networking and Applications (AINA), 2012, pp. 763–770, doi: 10.1109/AINA.2012.112.

[3] S. Gaur and V. Barthwal, "An extensive analysis of digital image watermarking techniques," Int. J. Intelligent Systems and Applications in Engineering, vol. 12, no. 1, pp. 121–145, Dec. 2023.

[4] M. S. Rana, M. M. Hasan, and S. K. S. Shuva, "Digital image watermarking using discrete wavelet transform and discrete cosine transform with noise identification," in Proc. 2nd Int. Conf. Intelligent Technologies (CONIT), 2022, pp. 1–4, doi: 10.1109/CONIT55038.2022.9847745.

[5] T. M. Thanh and K. Tanaka, "A novel q-DWT for blind and robust image watermarking," in Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC), 2014, pp. 2061–2065, doi: 10.1109/PIMRC.2014.7136511.

[6] T. M. Thanh and N. T. Thanh, "Extended DCT domain for improving the quality of watermarked images," in Proc. Int. Conf. Knowledge and Systems Engineering (KSE), 2015, pp. 336–339, doi: 10.1109/KSE.2015.70.

[7] M. Iwakiri and T. M. Thanh, "Fundamental incomplete cryptography method for digital rights management based on JPEG lossy compression," in Proc. IEEE Int. Conf. Advanced Information Networking and Applications (AINA), 2012, pp. 755–762, doi: 10.1109/AINA.2012.111.

[8] Y. Xia et al., "An adaptive blind color watermarking scheme based on Hadamard transform and information mapping system," Circuits, Systems, and Signal Processing, vol. 44, no. 5, pp. 3432–3465, Dec. 2025, doi: 10.1007/s00034-024-02971-0.

[9] A. S. Beggari, A. Wali, A. Khaldi, M. R. Kafi, and A. K. Sahu, "Secure and imperceptible medical image watermarking via multiscale QR embedding and attention-based optimization," *Engineering Science and Technology, an International Journal*, vol. 73, p. 102250, 2026, doi: 10.1016/j.jestch.2025.102250.

[10] A. K. Sahu and S. Mishra (eds.), *Fortressing Pixels: Information Security for Images, Videos, Audio and Beyond*. London, UK: IET, 2024, doi: 10.1049/PBSE030E.

[11] M. M. Sayah, N. Zermi, A. Khaldi, and M. R. Kafi, "ECG signal protection using redundant discrete wavelet transform-based data hiding," in *Fortressing Pixels: Information Security for Images, Videos, Audio and Beyond*, S. Deb, A. A.-A. Gutub, and A. K. Sahu, Eds. London, UK: The Institution of Engineering and Technology (IET), 2025, ch. 2, doi: 10.1049/PBSE030E_ch2

[12] Bekkari, F., Kafi, M. & Khaldi, A. Hybrid deep semantic query expansion using multi-objective fireworks-transformer optimization: a reinforcement learning approach. *Evol. Intel.* 18, 122 (2025). https://doi.org/10.1007/s12065-025-01109-8

[13] A. S. Beggari, A. Wali, A. Khaldi, M. R. Kafi, and A. K. Sahu, "Robust medical image watermarking based on Ridgelet transform and Ant Colony Optimization for telemedicine security," *Systems and Soft Computing*, vol. 7, art. no. 200390, 2025, doi: 10.1016/j.sasc.2025.200390.

[14] A. S. Beggari, A. Wali, A. Khaldi, M. R. Kafi, and A. K. Sahu, "Robust and imperceptible medical image watermarking for telemedicine applications based on transform-domain and neural clustering techniques," *Journal of the Franklin Institute*, vol. 362, no. 15, art. no. 108039, 2025, doi: 10.1016/j.jfranklin.2025.108039.

[15] A. S. Beggari, A. Wali, A. Khaldi, M. R. Kafi, and A. K. Sahu, "FDCT-based watermarking for robust and imperceptible medical image protection," *Intelligence-Based Medicine*, vol. 12, art. no. 100280, 2025, doi: 10.1016/j.ibmed.2025.100280.

[16] A. K. Sahu and M. Sahu, "Hybrid fragile image watermarking for tamper detection, localization and dual self-recovery," *Engineering Science and Technology, an International Journal*, vol. 73, art. no. 102266, 2026, doi: 10.1016/j.jestch.2025.102266.

[17] Y. Fang et al., "Robust zero-watermarking algorithm for medical images based on SIFT and bandelet-DCT," Multimedia Tools and Applications, vol. 81, no. 12, pp. 16863–16879, May 2022, doi: 10.1007/s11042-022-12592-x.

[18] Q. Su, D. Liu, and Y. Sun, "A robust adaptive blind color image watermarking scheme for resisting geometric attacks," Information Sciences, vol. 606, pp. 194–212, 2022, doi: 10.1016/j.ins.2022.05.046.

[19] H.-T. Hu and T.-T. Lee, “Robust complementary dual image watermarking in subbands derived from the Laplacian pyramid, discrete wavelet transform, and directional filter bank,” Circuits, Systems, and Signal Processing, vol. 41, no. 7, pp. 4090–4116, Jul. 2022, doi: 10.1007/s00034-022-01975-y.

[20] Y. Yuan et al., “Robust zero-watermarking algorithm based on discrete wavelet transform and daisy descriptors for encrypted medical images,” CAAI Transactions on Intelligence Technology, vol. 9, no. 1, pp. 40–53, Jan. 2024, doi: 10.1049/cit2.12282.

[21] X.-C. Yuan and M. Li, “Local multi-watermarking method based on robust and adaptive feature extraction,” Signal Processing, vol. 149, pp. 103–117, 2018, doi: 10.1016/j.sigpro.2018.03.007.

[22] T. M. Thanh et al., “Robust semi-blind video watermarking based on frame-patch matching,” AEU – Int. J. Electronics and Communications, vol. 68, no. 10, pp. 1007–1015, 2014, doi: 10.1016/j.aeue.2014.05.004.

[23] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, “KAZE features,” in Computer Vision – ECCV 2012, A. Fitzgibbon et al., Eds. Berlin, Heidelberg: Springer, 2012, pp. 214–227.

[24] V. Q. Pham et al., “Geometrically invariant object-based watermarking using SIFT features,” in Proc. IEEE Int. Conf. Image Processing (ICIP), vol. 5, 2007, pp. 473–476, doi: 10.1109/ICIP.2007.4379868.

[25] P. T. Hung and T. M. Thanh, “A strategy to select feature points for robust zero-watermarking algorithms,” in Proc. RIVF Int. Conf. Computing and Communication Technologies, 2024, pp. 63–67, doi: 10.1109/RIVF64335.2024.11009088.

[26] H.-H. Tsai, Y.-S. Lai, and S.-C. Lo, “A zero-watermark scheme with geometrical invariants using SVM and PSO against geometric attacks,” J. Systems and Software, vol. 86, no. 2, pp. 335–348, 2013, doi: 10.1016/j.jss.2012.08.040.

[27] S. A. Nawaz et al., “Medical image zero-watermarking algorithm based on dual-tree complex wavelet transform, AlexNet, and discrete cosine transform,” Applied Soft Computing, vol. 169, p. 112556, 2025, doi: 10.1016/j.asoc.2024.112556.

[28] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” Int. J. Computer Vision, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[29] C. E. Shannon, “A mathematical theory of communication,” Bell System Technical Journal, vol. 27, no. 3, pp. 379–423, Jul. 1948; vol. 27, no. 4, pp. 623–656, Oct. 1948.

[30] S. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[31] N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete cosine transform,” IEEE Trans. Computers, vol. C-23, no. 1, pp. 90–93, Jan. 1974.

[32] V. I. Arnold and A. Avez, Ergodic Problems of Classical Mechanics. New York, NY, USA: Benjamin, 1968.

[33] X. Liao, K. Wong, and S. Chen, “A novel image encryption algorithm based on chaotic maps,” Chaos, Solitons & Fractals, vol. 36, no. 2, pp. 432–444, 2008.

[34] N. K. Thanh et al., “A public-key encryption–authentication scheme based on the ElGamal cryptographic algorithm,” J. Science and Technique – Information and Communication Technology, vol. 12, no. 1, 2023, doi: 10.56651/lqdtu.jst.v12.n1.658.ict.

[35] Y. Zheng, “Digital signcryption or how to achieve cost (signature & encryption) ≪ cost (signature) + cost (encryption),” in Advances in Cryptology – CRYPTO’97, LNCS, vol. 1294. Berlin, Heidelberg: Springer, 1997, pp. 165–179.

[36] National Institute of Standards and Technology, Digital Signature Standard (DSS), FIPS PUB 186, U.S. Department of Commerce, May 1994.