Journal of Data Science and Modeling, Vol. 4, No. 1, 31-58, December 2025

Research Manuscript

A Hybrid Image Watermarking Method Based on Hessian Matrix Eigenvalues and Wavelet Transform for Digital Copyright Protection

Hadis Heidari*[1], Mosayyeb Amini[1]

[1] Kermanshah University of Technology, Kermanshah, Iran.

Received: 25/07/2026 Accepted: 09/11/2026

Abstract:

The widespread exchange of images over the internet and the growing threats of unauthorized copying and content manipulation have made digital image watermarking essential for protecting intellectual property and ensuring image authenticity. A major challenge in designing watermarking algorithms is achieving imperceptibility while preserving image quality. This paper proposes a hybrid digital image watermarking approach based on Hessian matrix eigenvalues and the discrete wavelet transform. After decomposing an image into frequency sub-bands using the wavelet transform, regions with stable structural and textural information are identified based on the Hessian matrix eigenvalues, and suitable coefficients for watermark embedding are selected. Combining spatial features derived from the Hessian matrix eigenvalues with the multi-scale analysis capability of the wavelet transform enables effective watermark distribution within the image. The proposed method is evaluated using peak signal-to-noise ratio, structural similarity index measure, and normalized correlation. Experimental results demonstrate that the method preserves watermark imperceptibility, making it suitable for security applications, digital copyright protection, and image authenticity verification.

Keywords: Digital image watermarking, Discrete wavelet transform (DWT), Hessian matrix eigenvalues, Copyright protection.
Mathematics Subject Classification (2010): 68U10, 94A08.

*Corresponding Author: hadis.heidari@kut.ac.ir

1. Introduction

The rapid growth of digital technologies, the widespread expansion of the internet, and the development of social networks have led to the massive generation, storage, and transmission of digital images within a short period. In addition to their widespread use in areas such as social media, digital images are also extensively employed in domains such as the protection of patient information in medical images, including magnetic resonance imaging and ultrasound (Elbasi (2020)), as well as military, legal, banking, intelligent surveillance, and authentication systems. However, digital images can be easily copied, duplicated, manipulated, and distributed without authorization. Consequently, the protection of intellectual property (Hafidz *et al.* (2025)), ensuring the authenticity of digital images, and preventing unauthorized modifications have become major challenges in the field of information security.

One effective approach to addressing these threats is digital image watermarking, in which information such as an owner identifier, serial number, or confidential data is imperceptibly embedded into an image in a manner that preserves the watermark's imperceptibility while allowing the watermark to be extracted even after certain modifications and attacks. In fact, digital watermarking is a technology that embeds specific information into digital products to protect copyright and integrity, prevent unauthorized copying, and track digital products, thereby playing a crucial role in ensuring the security of digital content (Dai (2022)). Accordingly, digital image watermarking can be employed in areas such as intellectual property protection, content management and identification, file archiving, and monitoring the distribution and dissemination of data (Rai *et al.* (2024)).

Watermarking algorithms are generally developed in the spatial and transform domains. Spatial-domain methods, such as least significant bit substitution, are simple to implement and offer high processing speed; however, they exhibit limited robustness against attacks such as compression, noise, and geometric transformations. In contrast, transform-domain methods embed the watermark into the coefficients of transforms such as the discrete wavelet transform (DWT), discrete cosine transform (DCT), and discrete sine transform (DST). Owing to their exploitation of the frequency characteristics of images, these methods generally demonstrate greater stability and robustness.

Among these transforms, DWT is one of the most widely used approaches in watermarking system design due to its multi-scale analysis capability, frequency information separation, and simultaneous preservation of spatial and frequency characteristics. DWT decomposes an image into four sub-bands, namely LL, LH, HL, and HH, each representing a different component of the image information. These sub-bands represent the main low-frequency information of the image, edge-

related information, and high-frequency texture and detail information. The appropriate selection of a sub-band for watermark embedding plays a significant role in establishing a balance between watermark imperceptibility and robustness against various attacks. Therefore, the intelligent selection of wavelet coefficients can improve the performance of watermarking systems to some extent.

On the other hand, the extraction of local structures, edges, corners, and intensity variations in digital images can be achieved using second-order derivatives, which are sensitive to sudden changes in intensity. In fact, second-order derivatives are capable of identifying regions with high structural stability. Consequently, second-order derivative-based operators, such as the Laplacian and Hessian matrix, are widely used in applications including edge detection, feature extraction, image segmentation, and computer vision. Despite the numerous advantages of these methods, the direct use of Hessian matrix eigenvalues in digital image watermarking has received comparatively limited attention, and considerable potential remains for developing more robust watermarking algorithms based on this information.

In recent years, numerous studies have been conducted to enhance the robustness and improve the performance of image watermarking systems. Many of these methods have attempted to overcome the limitations of conventional approaches by combining wavelet transforms with techniques such as the discrete cosine transform, Fourier transform, singular value decomposition (SVD), optimization algorithms, fuzzy logic, and deep learning methods. Although these approaches have demonstrated satisfactory performance in many cases, some involve high computational costs and considerable implementation complexity due to their complex structures and multi-stage computational procedures. Furthermore, in a substantial portion of previous studies, structural information and local image characteristics have not been systematically and purposefully incorporated into the process of selecting suitable locations for watermark embedding. This issue can affect the balance among watermark imperceptibility, image quality, and algorithm robustness. Therefore, the development of a method that exploits the multi-scale capabilities of the wavelet transform, uses the structural characteristics of the image to intelligently select suitable embedding regions, and maintains an acceptable level of computational complexity remains a significant research challenge.

In this paper, a digital image watermarking method based on the combination of Hessian matrix eigenvalues and the discrete wavelet transform is presented to address the existing limitations. In the proposed method (PM), the image is first decomposed into different frequency sub-bands using the wavelet transform. Subsequently, stable regions with suitable structural characteristics are identified

based on the features extracted from the Hessian matrix eigenvalues, and appropriate coefficients for watermark embedding are selected. This strategy enables the watermark information to be embedded in image regions that preserve the watermark's imperceptibility. To evaluate the performance of the proposed method, a series of experiments is conducted on different images. The quality of the watermarked images is evaluated using the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) metrics, while watermark robustness is assessed using the normalized correlation (NC) coefficient. The experimental results demonstrate that combining structural information derived from the eigenvalues of the Hessian matrix with the multi-scale analysis capability of the wavelet transform significantly enhances watermark imperceptibility while preserving the visual quality of the host image. The proposed method has the potential to be applied in various fields, including intellectual property protection, medical image authenticity verification, digital document security, intelligent surveillance, and secure digital image transmission.

The remainder of this paper is organized as follows. Section 2 presents a review of previous studies on various digital image watermarking methods. Section 3 describes the proposed image watermarking method based on the discrete wavelet transform and information extracted from Hessian matrix eigenvalues, along with its application areas. Section 4 evaluates the performance of the proposed method using various evaluation metrics, compares it with existing methods, and discusses the experimental results. Finally, Section 5 summarizes the main findings, discusses the limitations of the proposed method, and outlines directions for future research.

2. Review of Previous Research on Digital Image Watermarking

Digital image watermarking has become one of the most important research areas in multimedia security. The studies conducted to date can generally be classified into four main categories: spatial-domain methods, transform-domain methods, hybrid methods, and machine learning- and deep learning-based methods.

The first category mainly comprises image watermarking algorithms developed in the spatial domain, where watermark information is directly embedded into the pixel values of the image. The most well-known technique in this category is the least significant bit (LSB) method. Although LSB-based methods are widely used in many applications due to their simplicity, high processing speed, and low computational cost, they exhibit limited robustness against operations such as compression, noise addition, rotation, cropping, and resizing. Consequently, they

are not particularly suitable for advanced security applications.

The second category consists of transform-domain methods. Among the most widely used transforms is the discrete cosine transform, which enables the design of algorithms with improved resistance to compression. In these methods, the watermark is generally embedded into the middle-frequency coefficients to achieve an appropriate balance between image quality and robustness. The discrete wavelet transform, by decomposing an image into different frequency sub-bands, enables the selection of an appropriate region for watermark embedding and provides a desirable trade-off between image quality and resistance to attacks. Numerous studies have demonstrated that embedding watermarks in middle- or low-frequency sub-bands can improve the robustness of watermarking systems against compression, noise, filtering, and resizing attacks. The discrete Fourier transform (DFT) is also employed in applications where the probability of geometric attacks is relatively high. In addition, singular value decomposition exhibits stability against minor image variations. Therefore, embedding a watermark into singular values results in only a slight degradation in image quality while providing appropriate resistance against many attacks. For this reason, SVD has become one of the most widely used techniques in image watermarking.

In recent years, researchers have increasingly focused on the development of hybrid watermarking methods. These methods combine the advantages of multiple transforms to reduce the limitations associated with individual techniques. For example, DWT-DCT, DWT-SVD, DCT-SVD, and DWT-DCT-SVD combinations are among the most common hybrid architectures employed in various studies to achieve greater robustness, improved image quality, and higher embedding capacity. The results of previous studies indicate that these methods generally outperform algorithms based on a single transform.

In addition to classical methods, the development of machine learning and, particularly, deep learning has brought about a significant transformation in the field of watermarking. In these approaches, watermark embedding and extraction processes are learned by computational models rather than being manually designed. Convolutional neural networks (CNNs), autoencoders, generative adversarial networks (GANs), and transformer-based models, such as vision transformers, are among the architectures that have been employed in recent years to develop intelligent watermarking systems. These models are capable of automatically extracting robust image features and generating watermarks with improved quality and greater resistance to attacks.

In Tang *et al.* (2020), an image watermarking algorithm based on edge detection, compressed sensing theory, and singular value decomposition was proposed. In this method, the image was divided into non-overlapping blocks, and the most

appropriate block for information embedding was selected. The obtained results demonstrated that the proposed method achieved stable and satisfactory performance. In Pallaw *et al.* (2025), evolutionary algorithms, including the genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization, and firefly algorithm, were employed to improve the performance of the watermarking method. The application of these algorithms enhanced security and improved the ability to defend against unauthorized access, particularly in medical applications. In Srinivasan *et al.* (2024), a novel wavelet-based method for image watermarking was proposed, and its performance was evaluated on various image datasets from the Alaska2 Image Steganalysis dataset available on Kaggle. In Shivdeep *et al.* (2020), a security mechanism for medical applications based on reversible watermarking (RW) and prediction error expansion (PEE) was also presented. In this method, a private key was employed as a security parameter, and authentication was performed through watermarking. The proposed algorithm can be used to secure medical images, including ultrasound, magnetic resonance imaging (MRI), computed tomography (CT), and X-ray images. It also enables the embedding of information such as patient details, diagnostic summaries, or other relevant medical data into medical images. Within this framework, physicians or other authorized individuals can extract and identify the embedded information using the private key.

In Melman *et al.* (2023), a novel image watermarking algorithm based on the discrete cosine transform was proposed to achieve a balance between embedding imperceptibility and watermark robustness. In this method, pattern-based redundant embedding and synchronization blocks were employed to improve watermark robustness against image processing operations. Moreover, optimization of the patterns using a GA resulted in a high level of imperceptibility and satisfactory performance in terms of key evaluation metrics. In Bhavani *et al.* (2020), a hybrid algorithm based on image segmentation, singular value decomposition, discrete wavelet transform, homomorphic filtering, and RSA (Rivest-Shamir-Adleman) encryption was proposed to address various security threats. In this method, two segmented watermarks were generated from the original watermark using two distinct threshold values and embedded into the image. These threshold values were transmitted to the receiver as secret keys using the Diffie–Hellman key exchange algorithm and were employed in the copyright verification process. The performance of the proposed method was evaluated in the presence of Gaussian noise, and the obtained results demonstrated satisfactory performance.

In Duang *et al.* (2022), a discrete cosine transform-based method for digital watermarking was presented. In this method, watermark image loading, preprocessing, embedding, extraction, and the evaluation of watermark robustness

against various attacks were performed, and the embedded watermark information was extracted from image blocks. By employing a block-based DCT scheme, the presented algorithm demonstrated appropriate robustness and can be used to protect copyright-related information. In Shivdeep *et al.* (2020), the pixel values of the cover image were modified based on the similarity between the cover image and the watermark. In this method, only authorized users possessing the correct pseudonoise code were able to extract the watermark. Since the proposed method belongs to the category of invisible watermarking techniques, it does not significantly affect the appearance of the original image. In addition, hardware implementation of the algorithm was performed on FPGA (Field-Programmable Gate Array) and ASIC (Application-Specific Integrated Circuit) platforms.

In Harshavardhansai and Sheela (2024), a fuzzy logic-based watermarking method was compared with a hybrid DWT-DCT method. In Xie and Liu (2023), an efficient image watermarking algorithm was proposed to counteract tampering while minimizing distortion. Furthermore, an adaptive watermarking algorithm based on chaotic key segmentation was proposed to improve the quality of the watermarked image. In Yuan *et al.* (2025), a novel adaptive binary image watermarking algorithm with appropriate resistance to print–scan attacks was proposed. In this method, the binary image was first divided into separate blocks, and the number of black pixels in each block was counted. The binary image was then converted into a grayscale image, and the discrete cosine transform was applied to the resulting grayscale image. Subsequently, an adaptive optimization-based algorithm was proposed to modify the appropriate DCT coefficients for watermark embedding.

The watermark imperceptibility constraint was also considered to ensure image quality and optimize watermark strength. After coefficient modification, the transformed image was converted back to the spatial domain, and pixel modifications in each block of the binary image were determined. Finally, edge detection was employed to select and modify appropriate pixels in the binary cover image.

In Chotikawanid and Amornraksa (2023), a novel embedding component for spatial-domain image watermarking was proposed to protect copyright and verify ownership of digital images. In this method, a component known as the q-logarithm was extracted from the color host image and directly used for watermark embedding. During the extraction process, the original non-watermarked component was first estimated using the mean of the watermarked components within a small image region, thereby enabling blind watermark extraction. In addition, various parameter values were investigated to determine the optimal value of the parameter q that resulted in more accurate watermark extraction. The evaluation results demonstrated that the proposed method achieved superior performance in

terms of accuracy and robustness, even in the presence of attacks. In Hemanth *et al.* (2024), a novel enhanced embedding algorithm (EmA) and a lightweight encryption algorithm (LEA), together with a robust medical image watermarking method, were proposed.

In Tyagi and Singh (2026), the FedOwner framework was proposed to provide more effective protection for image classification models by combining black-box and white-box watermarking techniques. The framework is based on three main mechanisms: generating a unique trigger set (watermark) for each user using identity information hashing and a chaotic map; adding a watermarked adapter layer to verify ownership without reducing model accuracy; and identifying malicious users by evaluating the model with the user-specific trigger set and comparing prediction confidence scores. In Senapati *et al.* (2026), a blind, robust, and computationally efficient watermarking method for medical images was presented. In this method, the discrete Chebyshev transform, Hessenberg-SVD decomposition, and fruit fly optimization (FFO) algorithm were employed to optimally select embedding locations. To enhance security, the watermark was encrypted before embedding using an improved Hénon chaotic map. Unlike deep neural network-based methods, a lightweight denoising autoencoder (DAE) was used only during the extraction stage to improve the quality of the recovered watermark.

In Li *et al.* (2026), a zero-watermarking method for medical images was proposed to protect copyright, preserve image integrity, and improve resistance against complex attacks. In this method, the image was first encrypted using a two-dimensional logistic–sine chaotic map. Subsequently, the DINOv3-ViT network was employed to extract high-level semantic features, while Zernike moments (ZMs) were used to extract geometric and spatial features at different scales. Finally, the extracted features were combined with the copyright image using the XOR (exclusive OR) operator to generate a zero-watermark without modifying the medical image.

In Peng *et al.* (2026), the INNA (invertible neural network-based attack) framework was introduced as the first watermark removal method based on invertible neural networks. The objective of this method was to establish a balance among visual image quality, watermark removal effectiveness, and computational efficiency. Within this framework, the watermarked image is decomposed into two components: a low-resolution structural component and latent representations containing watermark information. A transformer-based module then selectively removes watermark information from the latent space, and the final image is reconstructed using the inverse mapping of the network. In Niu *et al.* (2026), a zero-watermarking framework for color images based on fast and accurate quaternion Chebyshev–Fourier moments (FAQCHFMs) was proposed to improve efficiency,

robustness, and security. The proposed method addresses three major limitations of moment-based approaches. First, a fast computational method for calculating fast Chebyshev–Fourier moments was developed. Subsequently, quaternion theory was employed to effectively extract the color features of the image. Finally, full four-dimensional (Full-4D) invariant features were combined using an asymmetric tent chaotic map to generate a zero-watermark.

In Jing (2020), a deep learning-based copy-resistant watermarking algorithm for color images was proposed. The obtained results demonstrated that the proposed method exhibited appropriate resistance to noise, filtering, and cropping attacks. In this method, the YUV color space was selected, and the chrominance component was binarized. In Badoga *et al.* (2021), a frequency-domain color image watermarking technique based on teaching–learning-based optimization (TLBO) and Lagrangian twin support vector regression (LTSVR) was presented. After single-level decomposition of the image, the statistical features of the selected wavelet band were analyzed, and LTSVR was employed in the watermark extraction and embedding processes. Moreover, TLBO was used to determine the optimal watermark strength for different selected blocks in the wavelet domain. Experimental results demonstrated that the proposed method exhibited greater resistance to JPEG (Joint Photographic Experts Group) compression than conventional methods.

In Liu *et al.* (2024), a robust watermarking algorithm for diffusion tensor images (DTIs) was developed by combining convolutional layers with a transformer. In this method, maximum correlation was employed as the most important feature weight to improve the quality of the reconstructed DTI image. In Wang *et al.* (2026), a deep image watermarking method based on a spatial–channel interaction attention generative adversarial network (SCIA-GAN) was proposed to improve watermark robustness against common processing operations on online social networks (OSNs). Unlike conventional methods that primarily embed watermarks in high-frequency components, the proposed method exploits semantic image features to achieve more stable watermark embedding. The proposed architecture includes a spatial–channel attention module, a message-guided processor for improving the distribution of watermark information, and a feature-preserving loss function. The use of semantic and wavelet features in this loss function improved the robustness of the proposed method against JPEG compression.

Although previous image watermarking methods have achieved satisfactory results in terms of imperceptibility and watermark extraction, several limitations can be identified from a methodological perspective. DWT-based methods provide an effective framework for embedding watermarks at different frequency levels; however, methods that select embedding coefficients mainly according to fixed fre-

quency bands may not sufficiently consider the local structural characteristics of the host image. Some approaches incorporate second-order derivative information, such as the Laplacian operator, to identify edge-rich or high-frequency regions for watermark embedding. While these methods can exploit areas with greater local variations, their selection strategies may not fully distinguish between structurally suitable and unsuitable regions. Other methods combine DWT with additional transforms or decomposition techniques to improve robustness, but this may increase computational complexity and introduce additional parameters that require careful adjustment.

Also, the emergence of technologies such as artificial intelligence-generated images, cloud-based multimedia systems, the Internet of Things (IoT), and social networks has introduced new challenges for researchers. Therefore, the development of intelligent, adaptive, and robust methods capable of resisting complex attacks remains one of the most important research directions in the field of digital image watermarking.

3. Proposed Digital Image Watermarking Method Based on the Discrete Wavelet Transform and Hessian Matrix Eigenvalues

In this paper, structural information derived from the eigenvalues of the Hessian matrix is intelligently combined with the discrete wavelet transform to develop a digital image watermarking method. Unlike methods that select wavelet coefficients randomly or according to fixed rules, the proposed method exploits structural image features to identify and determine appropriate embedding regions. Since regions with greater structural stability generally exhibit higher resistance to conventional image processing operations, the extracted watermark is expected to demonstrate greater stability.

The overall process of the proposed algorithm consists of six main stages, which are described below:

- Preprocessing of the host image and watermark image
- Image decomposition using the discrete wavelet transform
- Calculation of the Hessian matrix eigenvalues
- Identification of stable regions based on the eigenvalues
- Watermark embedding into the selected coefficients

- Watermark extraction and recovery

For image preprocessing, let the host image and watermark image be denoted by I(x,y) and W(x,y), respectively. First, the host image is converted into a grayscale image, and its intensity values are normalized within the range [0,255]. Subsequently, an initial denoising operation is performed using a weighted Gaussian filter to prevent random noise from affecting the feature extraction process. The Gaussian function is defined according to Eq. 3.1.

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\{-\frac{x^2+y^2}{2\sigma^2}\}. \tag{3.1}$$

The Gaussian filter within a (2k+1)×(2k+1) neighborhood is defined according to Eq. 3.2, where S is a normalization factor that ensures that the sum of the filter matrix elements equals one.

$$\frac{1}{S} \times \begin{pmatrix} G(-k,-k) & \dots & G(-k,0) & \dots & G(-k,+k) \\ \dots & \dots & \dots & \dots & \dots \\ G(+k,-k) & \dots & G(+k,0) & \dots & G(+k,+k) \end{pmatrix} \tag{3.2}$$

In the next stage, the image is decomposed using a single-level discrete wavelet transform. The discrete wavelet transform is an efficient approach for analyzing and processing signals, including digital images, audio signals, and data obtained from various sources. In this method, a signal is decomposed into a set of wavelets, which are short oscillations with different frequencies and time intervals. After applying the discrete wavelet transform, the image is divided into four sub-bands, namely LL, LH, HL, and HH, each of which contains different information about the structure and texture of the image. The LL sub-band represents information related to smooth regions and the low-frequency components of the image, whereas the LH sub-band provides information about horizontal edges, the HL sub-band represents vertical edges, and the HH sub-band contains high-frequency components.

In the proposed method, watermark embedding is performed in the middle-frequency LH and HL sub-bands and, when necessary, in a portion of the HH sub-band. The discrete wavelet transform has also been employed for image watermarking in Alsabaan *et al.* (2025).

The next step involves using the eigenvalues of the Hessian matrix to identify stable regions in the image. For each pixel, the second-order derivatives of the image intensity function in the horizontal and vertical directions are computed according to Eqs. 3.3 and 3.4.

$$I_{xx} = \frac{d^2 I}{dx^2}. \tag{3.3}$$

$$I_{yy} = \frac{d^2 I}{dy^2}. \tag{3.4}$$

In addition, the mixed partial derivative I_{xy} is calculated. Consider a 5×5 image I, together with the kernel functions along the x- and y-axes, as illustrated in Figure 1.

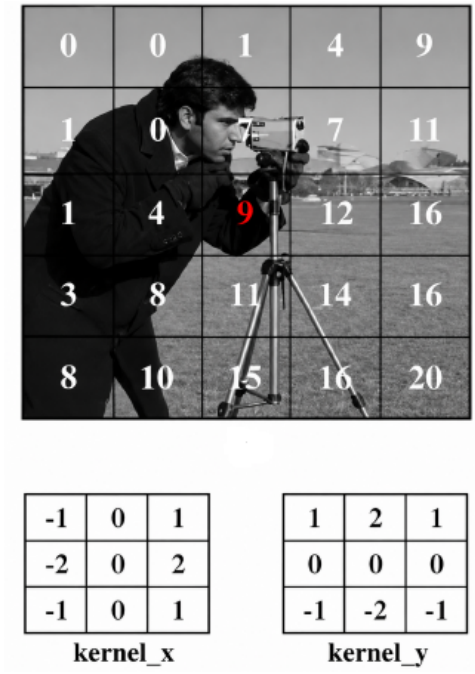


Figure 1: A 5×5 input image and kernel functions along the x- and y-directions.

For the neighborhood of the pixel located in the third row and third column, the first-order derivative along the x-axis is shown in Figure 2(a), while the first-order derivative along the y-axis is presented in Figure 2(b).

In general, suppose that the derivative of the input image along the x-axis, centered at I_c, has the values a_1, a_2, and a_3 in the first row; a_4, a_5, and a_6 in the second row; and a_7, a_8, and a_9 in the third row. Similarly, assume that the derivative of the input image along the y-axis has the values b_1, b_2, and b_3 in the first row; b_4, b_5, and b_6 in the second row; and b_7, b_8, and b_9 in the third row. The

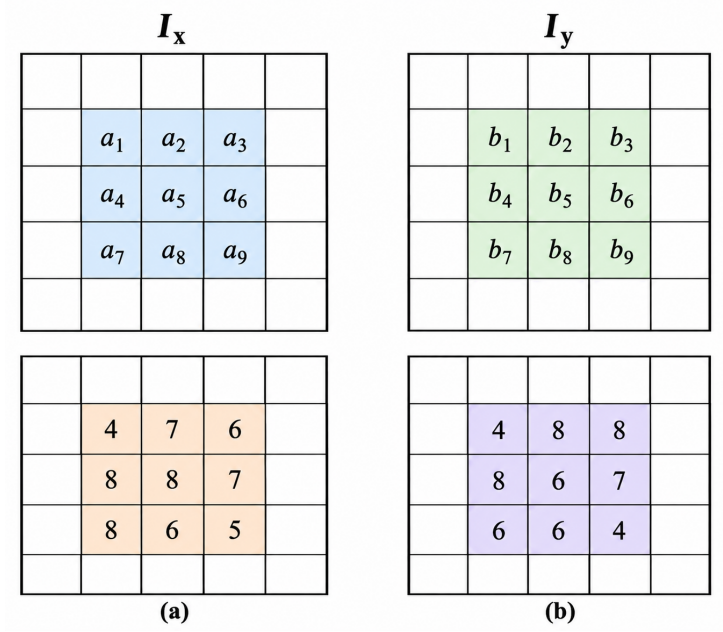


Figure 2: (a) First-order derivative along the x-axis in the neighborhood of the pixel located in the third row and third column, and (b) first-order derivative along the y-axis in the neighborhood of the same pixel.

product of the derivatives along the x- and y-axes is then obtained according to Eq. 3.5, yielding a value of 385 for this simple example.

$$I_x I_y = a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4 + a_5 b_5 + a_6 b_6 + a_7 b_7 + a_8 b_8 + a_9 b_9. \tag{3.5}$$

These derivatives form the Hessian matrix, as shown in Eq. 3.6.

$$H = \begin{pmatrix} I_{xx} & I_x I_y \\ I_x I_y & I_{yy} \end{pmatrix} = \begin{pmatrix} 403 & 385 \\ 385 & 381 \end{pmatrix} \tag{3.6}$$

For each point, the structural response is calculated based on the eigenvalues of the Hessian matrix. The structural response is obtained according to Eq. 3.7.

$$E = \frac{I_{xx} + I_{yy} + \sqrt{(I_{xx} - I_{yy})^2 + 4I_{xy}^2}}{2}. \tag{3.7}$$

This value represents the intensity of local variations in the image. Regions with values greater than the threshold T are selected as suitable regions for watermark embedding.

In fact, based on the eigenvalues of the Hessian matrix, candidate regions are evaluated according to their local structural characteristics, and those satisfying the predefined selection criterion and threshold are selected as suitable embedding regions for the subsequent watermark insertion process.

Unlike many existing methods that select wavelet coefficients in a fixed manner, the proposed method selects only those coefficients that simultaneously satisfy the following two criteria:

- Adequate wavelet energy
- A high second-order derivative response

This strategy ensures that the watermark is embedded in regions with stable image structures. One of the critical steps in enhancing the security of digital image watermarking is the use of effective encryption techniques to transform the message into an unreadable format.

The S-box is one of the mechanisms employed for this purpose (Tolpa *et al.* (2025)). Gaussian distribution, fractional linear transformation, the improved sine–cosine algorithm, fuzzy logic, and genetic algorithms are among the approaches that can be used for this purpose. Chaotic systems exhibit complex and unpredictable behavior and are therefore well suited for secure communications and cryptographic applications. The general mathematical model of a chaotic map is defined according to Eq. 3.8.

$$y_{m+1} = f(y_m). \tag{3.8}$$

In this equation, $f(y_m)$ denotes a function of y_m, where y_m represents the current state and y_{m+1} denotes the subsequent state of the system. The initial state of the map is set to y_0, while the output values are generated as the sequence y_1, y_2, ... In addition, the Lyapunov exponent (LE) of the discrete map is defined according to Eq. 3.9 (Tolpa *et al.* (2025)).

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)|. \tag{3.9}$$

In this equation, $f'(x_i)$ denotes the derivative of the function $f(x_i)$. For watermark embedding, the selected coefficients are denoted by C_i. The embedding process is performed according to Eq. 3.10.

$$C_i + \alpha W_i. \tag{3.10}$$

In this equation, C_i, W_i, and α denote the selected coefficient, the watermark bit, and the embedding strength factor, respectively. Although the S-box was not incorporated here, its application as an enhancement could potentially serve as an effective additional step in the proposed framework. Upon completion of the embedding process, the inverse wavelet transform is applied to generate the watermarked image. The proposed method offers the following advantages:

- Intelligent selection of embedding locations based on the structural characteristics of the image
- Preservation of visual image quality by avoiding modifications to sensitive coefficients
- Reduced watermark detectability
- Capability of implementation in real-time systems
- Achieving an appropriate balance between imperceptibility and embedding capacity

The proposed watermarking scheme employs a blind watermark extraction procedure. After watermark embedding, the watermark can be recovered directly from the watermarked image using the predefined extraction procedure and the required embedding information. This blind extraction characteristic improves the practical applicability of the proposed method, particularly in scenarios where the original image is unavailable at the time of watermark verification or recovery. Also, the embedding capacity of the proposed method is 4096 bits per host image.

The combination of structural information derived from the eigenvalues of the Hessian matrix with the multi-scale analysis capability of the wavelet transform enables the optimal selection of embedding coefficients. Consequently, the watermark is embedded in stable regions of the image.

For watermark extraction, the received watermarked image is first acquired. Subsequently, a single-level discrete wavelet transform is applied to the image, decomposing it into different sub-bands. The intermediate-frequency sub-bands are then selected, and the watermark is recovered from the HL and LH sub-bands. Finally, the recovered watermark blocks are combined to reconstruct the final watermark image.

The overall workflow of the proposed method is illustrated in Figure 3.

4. Experimental Results

To evaluate the performance of the proposed algorithm, a series of experiments was conducted on different images under various conditions. The purpose of this

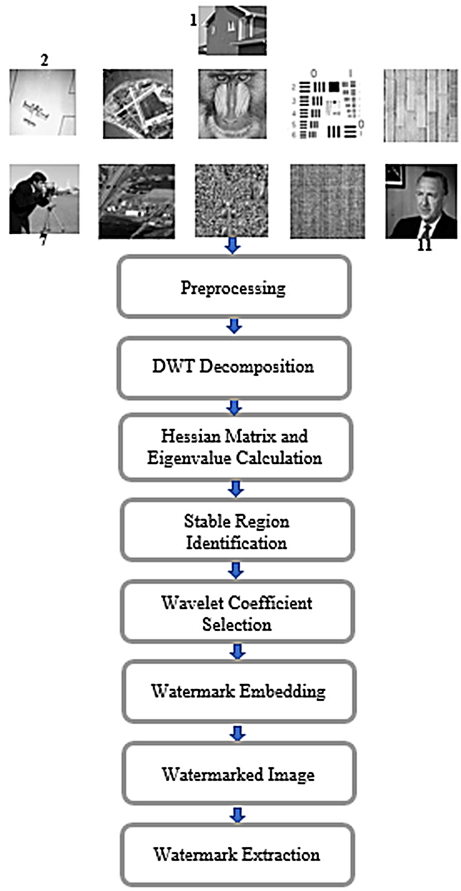


Figure 3: The overall procedure of the proposed method.

evaluation was to assess the performance of the proposed method in terms of watermarked image quality and watermark extraction accuracy. To comprehensively evaluate the algorithm, three standard metrics were employed. The quality of the watermarked image relative to the original image was evaluated using the peak signal-to-noise ratio, which was calculated according to Eq. 4.11.

$$PSNR = 10 \log \frac{I_{max}^2}{MSE}. \tag{4.11}$$

$$MSE = \frac{1}{n} \sum_{x} [I(x) - \widehat{I}(x)]^2. \tag{4.12}$$

In this equation, I_{max} denotes the maximum image intensity value, which is

equal to 255, while MSE represents the mean squared error. Higher PSNR values indicate better visual image quality. In general, PSNR values greater than 40 dB are considered indicative of highly desirable watermark imperceptibility.

The structural similarity index measures the structural similarity between the original and watermarked images and is calculated according to Eq. 4.13.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C2)}{(\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)}. \tag{4.13}$$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{4.14}$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2. \tag{4.15}$$

$$\sigma_{xy}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}). \tag{4.16}$$

Here, C_1 and C_2 are small numerical constants used to avoid numerical instability during the computation, and μ_x denotes the mean intensity value of the image.

The closer the SSIM value is to 1, the greater the structural similarity and, consequently, the higher the quality of the watermarked image.

To measure the similarity between the original and extracted watermarks, the normalized correlation metric is employed according to Eq. 4.17.

$$NC = \frac{\sum_i \sum_j W(i,j)W'(i,j)}{\sqrt{\sum_i \sum_j W^2(i,j)}\sqrt{\sum_i \sum_j W'^2(i,j)}}. \tag{4.17}$$

In this equation, W(i,j) and $W'(i,j)$ denote the original watermark and the extracted watermark images, respectively. The value of this metric ranges from 0 to 1, where NC=1 indicates completely accurate watermark extraction.

Figure 4 presents the original image, the watermark, the watermarked image, and the extracted watermark obtained using the proposed method. The PSNR, SSIM, and NC values were 37.97 dB, 0.9799, and 0.9324, respectively.

To ensure the reproducibility of the proposed method, all implementation parameters used in the experiments were fixed and explicitly specified. A single-level

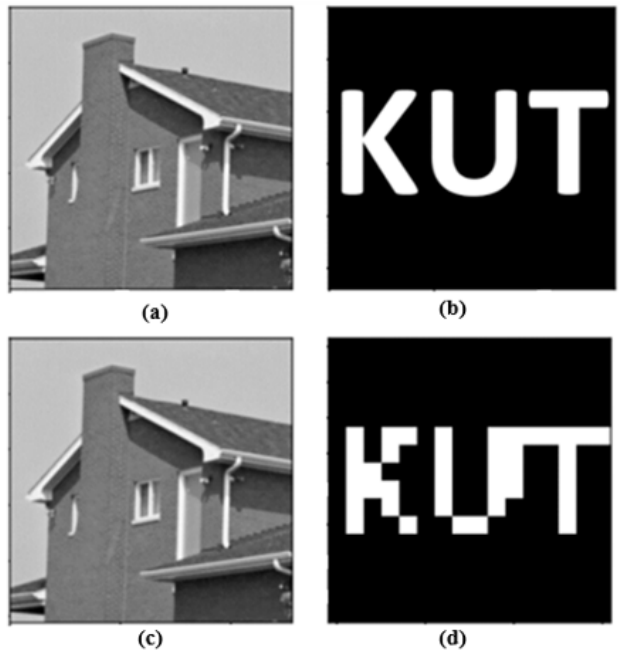


Figure 4: (a) Original image, (b) watermark, (c) watermarked image, and (d) extracted watermark obtained using the image watermarking method.

discrete wavelet transform was applied. For Gaussian filtering, a neighborhood size of k=3 and a standard deviation of σ=0.5 were employed. The threshold for stable-region selection was set to T=0.05. In all experiments, the host images were resized to 256×256 pixels, and the watermark images were set to 64×64 pixels. The same parameter settings were consistently applied to all test images to ensure a fair and reproducible evaluation of the proposed watermarking method.

Figure 5 presents additional images in which the selected watermark of the proposed method, introduced in Figure 3, has been embedded. The PSNR, SSIM, and NC values for the first image were 39.38 dB, 0.9891, and 0.9324, respectively. For the second image, the corresponding values were 28.97 dB, 0.9039, and 0.9324, respectively, while for the third image, they were 24.94 dB, 0.8406, and 0.9324, respectively. Moreover, the watermark extraction accuracy exceeded 99% for all three images. These results demonstrate the satisfactory performance of the proposed method and its applicability in various practical domains, including medical applications for embedding patient information.

Furthermore, the four images shown in Figure 6 correspond to cases in which the watermark extraction accuracies were 96.88%, 97.27%, 97.66%, and 98.83%, respectively. For the first image, the PSNR, SSIM, and NC values were 29.09 dB,

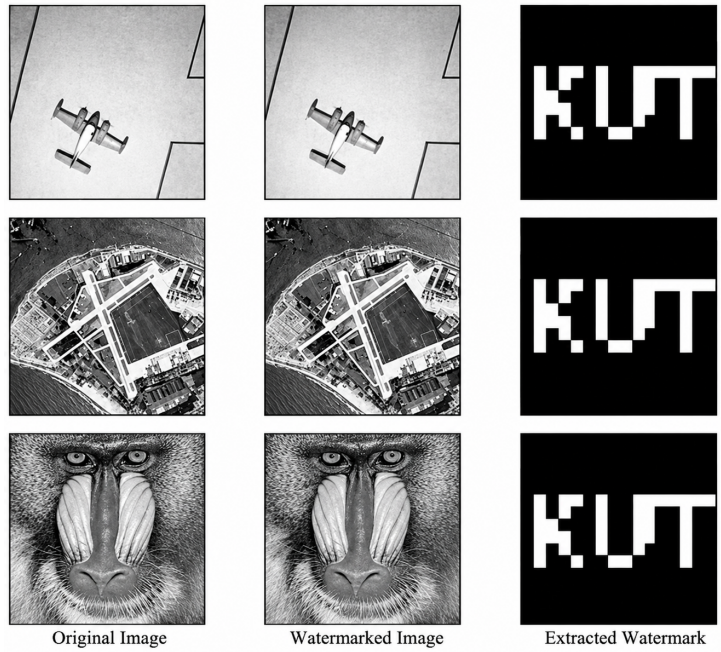


Figure 5: Original images, watermarked images, and watermarks extracted from the images.

0.9647, and 0.8733, respectively. For the second image, the corresponding values were 29.93 dB, 0.8702, and 0.8669, respectively. For the third image, these values were 28.97 dB, 0.8608, and 0.8890, respectively, whereas for the fourth image, they were 35.68 dB, 0.9701, and 0.9210, respectively. These results confirm the satisfactory performance of the proposed method.

In addition, Figure 7 presents three images in which the watermark has been embedded in such a manner that no perceptible difference can be observed between the watermarked and original images. The SSIM values for these images are 0.8579, 0.8819, and 0.9716, respectively. In this figure, the first row contains the original images, whereas the second row presents the corresponding watermarked images.

The experimental results demonstrate that the eigenvalues of the Hessian matrix can effectively identify regions containing high levels of detail and structural energy. Embedding the watermark in these regions makes the introduced modifications less perceptible to the human visual system. On the other hand, the wavelet transform decomposes the image into different frequency sub-bands, thereby enabling the selection of suitable coefficients for watermark embedding. The combination of these two tools enables the proposed algorithm to achieve an appropriate balance between imperceptibility and embedding capacity. Furthermore, adaptive

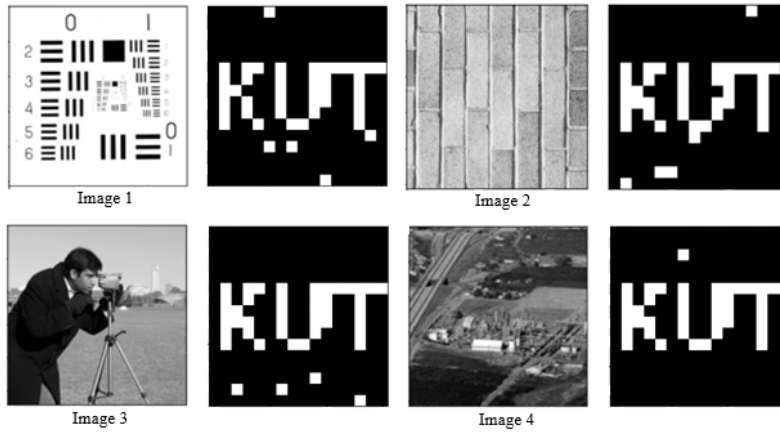


Figure 6: Watermarked images and their extracted watermarks for four different images.

adjustment of the embedding strength based on the second-order derivative response allows the watermark to be embedded with greater strength in regions with complex textures while maintaining minimal modifications in homogeneous regions. This characteristic contributes significantly to improving image quality and reducing the likelihood of watermark detection.

Despite the promising performance of the proposed algorithm, several limitations remain. Its robustness may decrease when subjected to severe geometric attacks, such as large-angle rotations, extensive cropping, or combinations of multiple consecutive attacks. Furthermore, the selection of optimal embedding strength parameters has a significant impact on both the imperceptibility of the watermark in the watermarked image and the robustness of the algorithm, requiring careful adjustment according to the image characteristics and the intended application.

The experimental results demonstrate that the proposed watermarking method, based on the second-order derivative and the discrete wavelet transform, achieves excellent performance in terms of image quality and watermark extraction accuracy. High PSNR and SSIM values indicate that the watermark imperceptibility and structural similarity of the watermarked image are well preserved. These findings indicate that the proposed method is a promising approach for applications involving digital copyright protection, image authenticity verification, and digital image security.

A comparison with the method proposed by Alsabaan et al. highlights several methodological differences between the two approaches. Alsabaan et al. use a Laplacian-based second-order derivative measure to identify sharp or edge-rich regions in the mid-frequency wavelet sub-bands. In contrast, the proposed method

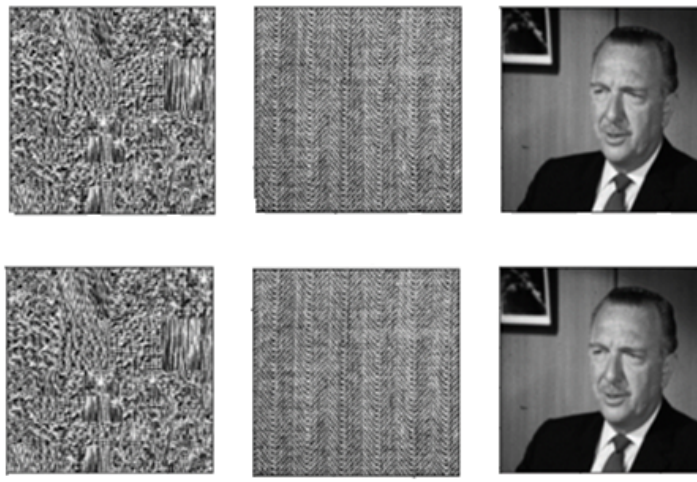


Figure 7: Watermark embedding in images using the proposed method.

employs a Hessian-based structural criterion in combination with wavelet-energy information to identify structurally suitable embedding regions and select the corresponding wavelet coefficients. In addition, the proposed framework incorporates a preprocessing stage before the DWT decomposition and structural analysis, which is not included in the embedding framework of Alsabaan et al. Furthermore, while the chaotic mechanism in Alsabaan et al. is mainly employed for watermark encryption, the proposed method focuses on adaptive region and coefficient selection based on structural and wavelet-domain characteristics.

The proposed method incorporates Hessian-based structural analysis to identify suitable embedding regions based on local image characteristics. This strategy introduces structural information into the region-selection process rather than relying solely on conventional transform-domain processing. Although the results reported by Alsabaan et al. provide useful context for evaluating watermarking performance, differences in datasets, experimental settings, implementation details, and evaluation criteria must be considered when interpreting results across the two studies. Therefore, a definitive quantitative comparison requires sufficiently consistent experimental conditions. The present study focuses on evaluating the proposed framework within its specified experimental setup and examining the contribution of Hessian-based region selection to the overall watermarking process.

Table 1 reports all the evaluation metrics obtained for the eleven test images, which are numbered sequentially in Figure 3. As can be observed, the proposed method achieves satisfactory performance.

Table 1: Evaluation results of the proposed method for the eleven test images.

Image Number	**PSNR (dB)**	**SSIM**	**NC**	**Accuracy (%)**
1	37.97	0.9799	0.9324	99.22
2	39.38	0.9891	0.9324	95.20
3	28.97	0.9039	0.9324	99.22
4	24.94	0.8406	0.9324	99.22
5	29.09	0.9647	0.8733	96.88
6	29.93	0.8702	0.8669	97.27
7	28.97	0.8608	0.8890	97.66
8	35.68	0.9701	0.9210	98.83
9	18.61	0.8579	0.6160	84.38
10	18.18	0.8819	0.7256	91.41
11	41.37	0.9716	0.7566	91.80

It is worth noting that relatively lower PSNR values were obtained for some host images, particularly 28.97 dB and 24.94 dB. This variation can be attributed to differences in the structural and textural characteristics of the host images. Since the proposed method adaptively identifies suitable regions and wavelet coefficients according to the structural information of each image, the number and location of modified coefficients may vary across different images. Images containing stronger edges, textures, or greater local intensity variations may therefore experience larger pixel-level modifications, resulting in lower PSNR values. Nevertheless, these values should be considered together with the SSIM and NC results, which provide complementary measures of visual similarity and watermark extraction accuracy. Therefore, the relatively lower PSNR values observed for these images reflect the content-dependent behavior of the proposed adaptive embedding strategy rather than a failure of the watermarking method.

Table 2 compares the average results obtained using the PM with those obtained using the basic discrete wavelet transform, Haar wavelet transform–discrete cosine transform (HWT-DCT), and discrete wavelet transform–discrete cosine transform (DWT-DCT) methods. The obtained results demonstrate the superior performance of the proposed method compared with the other methods considered.

As shown in Table 2, the proposed watermarking method achieves superior overall performance compared with the conventional DWT, HWT-DCT, and DWT-DCT approaches. This improvement can be attributed to the Hessian-based region selection strategy, which identifies structurally significant regions of the image for watermark embedding rather than relying solely on conventional transform-domain

Table 2: Comparison of the Evaluation Metrics of the PM with Existing Methods.

Methods	**PSNR (dB)**	**SSIM**	**NC**	**Average Accuracy (%)**
PM	30.28	0.9173	0.8525	**95.55**
DWT	27.14	0.8819	0.8387	93.60
HWT-DCT	28.18	0.8954	0.8402	94.91
DWT-DCT	29.32	0.9053	0.8499	95.03

coefficients. By exploiting local image structures and selecting more suitable embedding regions, the proposed method provides a more favorable balance between watermark imperceptibility and reliable watermark recovery. The comparative results therefore demonstrate the effectiveness of incorporating Hessian-based structural information into the watermarking framework and highlight the advantage of the proposed approach over the conventional and hybrid transform-based methods considered.

This study does not include a comprehensive parameter sensitivity analysis or a statistical evaluation based on multiple independent runs. Further evaluation will therefore examine parameter sensitivity and report statistical measures, such as the mean and standard deviation, across multiple experiments.

In this paper, no claims are made regarding the resistance of the proposed method to various attacks, as the present study did not include dedicated and systematic experiments to evaluate its performance under attacks such as JPEG compression, noise, filtering, cropping, rotation, and scaling. Therefore, based on the current results, no definitive conclusion can be drawn regarding the robustness or stability of the proposed method against these attacks. Further evaluation is required to comprehensively assess the robustness of the proposed method against various types of attacks.

5. Conclusion and Future Work

This paper proposed a digital image watermarking method based on the eigenvalues of the Hessian matrix and the discrete wavelet transform. By utilizing second-order derivatives to identify structurally significant regions, the watermark was adaptively embedded into appropriate wavelet coefficients. The results demonstrated the effectiveness of the proposed approach for applications such as copyright protection, authenticity verification, digital rights management, medical imaging, and secure multimedia systems.

Despite the promising results obtained by the proposed watermarking method,

some limitations should be acknowledged. First, the performance of the method may vary across images with different levels of complexity.

Second, the current experimental evaluation is based on a limited set of host images, and therefore, the generalizability of the proposed method should be further investigated using larger and more diverse image datasets. In addition, although the proposed method provides satisfactory results under the evaluated conditions, its robustness against a broader range of intentional and unintentional attacks requires further investigation.

An important limitation of this study is that the robustness of the proposed watermarking method has not yet been quantitatively evaluated against common image processing and geometric attacks. Although the experimental results provide an assessment of imperceptibility and basic watermark recovery, they do not establish the method's resistance to JPEG compression, additive noise, filtering, cropping, rotation, or scaling. Further investigation will focus on this limitation through systematic experiments using objective measures, including normalized correlation and bit error rate, together with comparisons against existing watermarking methods under various attack conditions. These investigations will provide a more comprehensive assessment of the trade-off between imperceptibility, watermark recovery, and robustness.

Further research may focus on evaluating the proposed method using larger and more diverse datasets, conducting comprehensive robustness experiments under various attack scenarios, and investigating optimization strategies for automatically determining the embedding parameters to achieve a better balance between imperceptibility and robustness. Additional investigations could explore ways to improve the trade-off among robustness, security, embedding capacity, and computational complexity through the integration of deep learning and intelligent optimization techniques. Other promising directions include enhancing resistance to geometric transformations and advanced attacks, extending the method to color images and video, and exploring its integration with cryptographic techniques, blockchain technology, and digital rights management systems.

Furthermore, the potential application of a chaotic S-box warrants further investigation.

References

Alsabaan M., Faheem Z. B., Zhu Y., and Ali J. (2025), Image watermarking algorithm based on the second-order derivative and discrete wavelet transform, *Computers, Materials & Continua,* **84**(1), 491–512.

Badoga N. K., Goyal R. K., and Mehta R. (2021), A color image watermarking in the frequency domain using a teaching-learning optimization algorithm,

6^{th} *International Conference on Image Information Processing (ICIIP),* Shimla, India, 1–5.

Bhavani Y., Puppala S. S., Pabba S., Kasarla K. S., and Anvitha K. (2020), Image segmentation-based hybrid watermarking algorithm for copyright protection, 11^{th} *International Conference on Computing, Communication and Networking Technologies (ICCCNT),* Kharagpur, India, 1–5.

Chotikawanid P. and Amornraksa T. (2023), Image watermarking algorithm based on q-algorithm component, *International Conference on Computer Graphics and Image Processing (CGIP),* Tokyo, Japan, 1–5.

Dai C. (2022), Analysis of digital watermarking technology and its applications, *International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI),* Zakopane, Poland, 1–5.

Duang Y., Wang Y., Cao C., and Zhang X. (2022), Research on digital watermarking algorithm based on discrete cosine transform, 15^{th} *International Congress on Image and Signal Processing,* Beijing, China, 1–5.

Elbasi E. (2020), Enhancing the robustness of watermarked medical images using a heuristic search algorithm, 7^{th} *International Conference on Internet of Things: Systems, Management and Security (IOTSMS),* Paris, France, 1–5.

Hafidz M. A., Ernawan F., Fakhreldin M., Adam Y. A., and Elhadi H. (2025), Robust image watermarking scheme based on a prediction error algorithm, 4^{th} *International Conference on Computer Technologies (ICCTech),* Kuala Lumpur, Malaysia, 1–5.

Harshavardhansai B. and Sheela D. (2024), Comparison of digital image watermarking using a hybrid DWT-DCT algorithm with a fuzzy-based algorithm for improving PSNR, *International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE),* Bangalore, India, 1–5.

Hemanth P., Bharathi P. S., and Sajiv G. (2024), Robust medical image watermarking scheme with a novel modified embedding algorithm in comparison to a lightweight encryption algorithm, 9^{th} *International Conference on Science Technology Engineering and Mathematics (ICONSTEM),* Chennai, India, 1–5.

Jing W. (2020), Research on digital image copying watermarking algorithm based on deep learning, *International Conference on Robots & Intelligent System (ICRIS),* Sanya, China, 1–5.

Li Y., Jiao M., and Xiong X. (2026), Robust zero-watermarking method for medical images via DINOv3-ViT and multi-scale neighborhood Zernike moments, *Neurocomputing,* **699**, 1–5.

Liu C., Li Z., Wang G., and Zheng L. (2024), A robust tensor watermarking algorithm for diffusion-tensor images, *Frontiers of Information Technology & Electronic Engineering,* **25**, 384–397.

Melman A., Eysutin O., and Smirnov D. (2023), An image watermarking algorithm in the DCT domain based on optimal patterns, *International Symposium Problems of Redundancy in Information and Control Systems,* Moscow, Russian Federation, 1–5.

Niu P., Wang H., Wang D., and Zhang Y. (2026), Efficient and accurate quaternion Chebyshev-Fourier moments for color image zero-watermarking, *Journal of the Franklin Institute,* **363**, 1–5.

Pallaw V. K., Kumar P., Pandey N. K., and Singh S. K. (2025), Enhancing security in medical imaging: optimized nature-inspired algorithms for robust watermarking in grayscale images, *International Conference on Pervasive Computational Techniques (ICPCT),* Greater Noida, India, 1–5.

Peng Y., Wang N., Qin C., Feng G., and Zhang X. (2026), INNA: An invertible neural network-based attack framework for robust image watermarking, *Information Fusion,* **135**, 1–5.

Rai A., Singh S., and Kushwaha S. (2024), Hybrid watermarking techniques in medical imaging: a comprehensive analysis and performance evaluation, *International Conference on Communication, Control, and Intelligent Systems (CCIS),* Mathura, India, 1–5.

Senapati R. K., Madhu G., Kautish S., Swain G., Mohamed A. W., and Almazyad A. (2026), A moment-preserving transform-domain robust medical image watermarking with extraction-side learning, *International Journal of Electronics and Communications,* **216**, 1–5.

Shivdeep, Ghosh S., Ghosal P., Maity S. P., and Rahaman H. (2020), PEE-based reversible watermarking algorithm for authentication and security of medical images, *IEEE Region 10 Symposium (TENSYMP),* Dhaka, Bangladesh, 1–5.

Shivdeep, Ghosh S., and Rahaman H. (2020), A new digital color image watermarking algorithm with its FPGA and ASIC implementation, *International Symposium on Devices, Circuits and Systems (ISDCS),* Howrah, India, 1–5.

Srinivasan R., Khilar R., and Haider W. (2024), Increasing the peak signal-to-noise ratio in encoding medical data with a cover image using a watermarking algorithm compared with a wavelet algorithm, 9^{th} *International Conference on Engineering Techniques and Applied Sciences (ICETAS),* Bahrain, 1–5.

Tang P., Zhang C., Pan W., Long T., and Zhang J. (2020), Edge detection and compressive sensing-oriented digital image block watermarking, *Chinese Automation Congress (CAC),* Shanghai, China, 1–5.

Tolpa S. H., Abdelhamed M. A., Said E., and Afi M. (2025), A novel chaos-based approach for constructing lightweight S-boxes, *Nature,* 1–5.

Tyagi T. and Singh A. K. (2026), FedOwner: Protecting image classification models using black-box and white-box watermarking, *Image and Vision Computing,* 1–5.

Wang R., Gu L., Li Y., Wang J., Wan W., Sun J., and Cheung S. S. (2026), SCIA-GAN: Robust image watermarking via spatial-channel interaction attention and feature preservation, *Expert Systems with Applications,* **328**, 1–5.

Xie R. and Liu H. (2023), An efficient anti-tampering image watermarking algorithm with low distortion, *International Conference on Image Processing and Computer Vision (IPCV),* Okinawa, Japan, 1–5.

Yuan Z., Li X., and Liu M. (2025), A novel print-scan-resilient binary image watermarking method based on adaptive DCT coefficient modification, 10^{th} *International Conference on Cloud Computing and Big Data Analytics (ICCCBDA),* Chengdu, China, 1–5.