Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 28 (2025), No. 8, pp. 3049–3058
DOI : 10.47974/JDMSC-2466

Analysis of the robustness and cryptographic security of a DCT-based image watermarking scheme

Govind Murari Upadhyay [#]
Department of Computer Applications
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Naveen Tewari [†]
School of Computing
Graphic Era Hill University
Bhimtal Campus
Nainital 263136
Uttrakhand
India

Surabhi Shanker [§]
Centre of Excellence-Cyber Security, School of Engineering and Technology
K. R. Mangalam University
Gurugram 122103
Haryana
India

Ajay Kumar Phogat [‡]
Department of Computer Science
Maharaja Surajmal Institute
Janakpuri 110058
Delhi
India

[#] *E-mail:* govind.upadhyay@jaipur.manipal.edu
[†] *E-mail:* navtewari@gmail.com
[§] *E-mail:* surabhi.shanker@krmangalam.edu.in
[‡] *E-mail:* ajayphogat@msijanakpuri.com

Aparna Misra [@]
Department of Computer Science & Information Technology
Trinity Institute of Professional Studies
Dwarka 110077
New Delhi
India

Timothy Malche *
Department of Computer Applications
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Abstract

Image watermarking is a crucial technique used to protect digital media from unauthorized access and ensure copyright security. In cryptography, image watermarking is the process of encoding visible or hidden data onto digital photographs for uses like secure communication, copyright protection, and authentication. By identifying unlawful changes, it protects image integrity and aids in ownership verification. However, the effectiveness of a watermarking system depends significantly on its robustness against various attacks and understanding how different attacks can affect the quality of watermarked images. In this paper, the performance of a discrete cosine transforms (DCT) based image watermarking scheme is analyzed in the presence of different security attacks like salt-pepper noise, median and motion attacks. The performance is evaluated both visually and by MSE parameters.

Subject Classification: *Primary 93A30, Secondary 49K15.*

Keywords: *Digital image watermarking security, Cryptography, DCT, Salt and pepper noise, Median attack, Motion attack, Industry innovation and infrastructure.*

1. Introduction

Digital image watermarking is a widely used technique for embedding information into digital images to ensure intellectual property protection, maintain data authenticity, and enable copyright management. Watermarking plays a significant role in various fields, such as copyright protection, digital forensics, authentication, and media monitoring. In an increasingly digital environment, digital image watermarking has become an essential tool for safeguarding intellectual property and maintaining

[@] *E-mail:* **aparna@tips.edu.in**
* *E-mail:* **timothy.malche@jaipur.manipal.edu (Corresponding Author)**

the integrity of digital assets. Strong watermarking techniques are crucial for protecting ownership rights and preserving the authenticity of digital media because of the abundance of multimedia content and the simplicity with which unauthorized parties can alter it. The robustness against attacks and imperceptibility of traditional watermarking techniques are frequent problems. As the demand for secure distribution and usage of digital content grows, watermarking has become a critical tool for addressing issues like unauthorized duplication and content modification. Despite its utility, digital image watermarking faces challenges that require ongoing research and innovation. Robustness is a key concern, as watermarks must withstand attacks such as noise addition, cropping, and compression. Maintaining imperceptibility is another challenge, as the watermark should not degrade the visual quality of the host image. Balancing embedding capacity with robustness and imperceptibility also poses difficulties. Furthermore, real-time implementation demands algorithms that are both efficient and computationally viable. The absence of standardized evaluation metrics complicates consistent benchmarking and comparison of watermarking techniques.

Potential attacks on any image watermarking scheme can be categorized into compression, geometric, noise and filtering attacks. Compression attacks distort the embedded watermark during image compression and are more effective at high-frequency components. Whereas geometric attacks involve transformation like rotation, scaling, translation etc can distort the spatial relationship between the watermark and image. The noise attacks arise due to the addition of noise like salt and pepper that can degrade the image quality and obscure the embedded watermark. The noise interferes with the watermark signal, and it is very hard to extract the embedded watermark under such circumstances. The filtering attack arises from filters like low, high and median filters that can significantly affect watermark visibility and detection by either smoothing or enhancing certain aspects of the image. A robust watermarking method should resist the effect of these attacks by embedding the watermark into a resilient component of the image and by analyzing these effects we can ensure the security and authenticity of the watermarking technique.

To address the issues and their effects on image watermarking, in this paper, the performance of a DCT-based image watermarking system under noise and different types of attack has been analyzed. The remaining paper is structured as follows, the related work is discussed in Section 2, the proposed methodology is presented in Section 3, and the findings are summarized in Section 4. Finally, the work is concluded in Section 5.

2. Related Work

Digital image watermarking has gained importance due to the surge in digital media usage and the growing necessity to safeguard content from unauthorized access. Researchers have developed various techniques, broadly categorized based on their domain of operation, robustness, and visibility.

Spatial domain techniques involve direct modification of pixel values in the image. A notable example is the Least Significant Bit (LSB) method, which embeds watermark information into the least significant bits of pixel values [1]. While computationally efficient, it is vulnerable to attacks such as compression and cropping. Another approach, the patchwork technique, alters pixel pairs using statistical methods, offering simplicity but limited resistance to geometric distortions [2]. In contrast, frequency-domain techniques transform the image into a frequency domain before embedding the watermark, resulting in improved robustness. The Discrete Cosine Transform (DCT) technique embeds the watermark in mid-frequency coefficients to balance imperceptibility and durability [3]. The Discrete Wavelet Transform (DWT) facilitates multi-resolution analysis, making it effective against compression and scaling [4]. Similarly, the Discrete Fourier Transform (DFT) and integration of DWT and DCT enhance robustness while preserving image quality. offers robustness to rotation and scaling attacks but requires significant computational resources [5]. Hybrid techniques, which combine spatial and frequency domain methods, exploit the advantages of both approaches. Rai and Singh [6] introduce a robust hybrid watermarking technique for securing medical images using SVM that identifies and preserves critical diagnostic areas by separating Regions of Interest (ROI) from non-ROI during watermark embedding. The method achieves high image quality, maintaining diagnostic integrity through metrics like PSNR and SSIM. Similarly, Lagrangian Support Vector Regression (LSVR) in the wavelet domain enable effective watermark embedding while preserving the quality of the image and improving resilience against various attacks [7]. The experimental results show that the method achieves a good balance between robustness and image quality, making it suitable for secure watermarking applications. Recent advancements in digital image watermarking include the integration of machine learning and blockchain technologies. Deep neural networks have been utilized for adaptive watermarking, achieving significant improvements in robustness and imperceptibility [8]. A deep learning-based image watermarking method

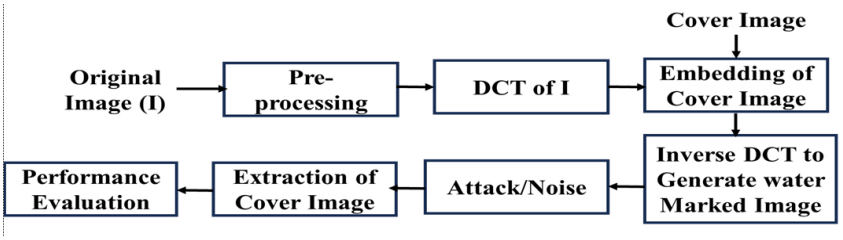

Figure 1

Flow chart of the proposed method

is designed to counter attacks and emphasizes the importance of designing watermarking systems with integrated security measures to safeguard digital content from unauthorized interference and explores the use of Fully Convolutional Neural Networks (FCNN) as a denoising attack, which removes watermarks while preserving image quality. By evaluating watermarking schemes like Spread Transform Dither Modulation (STDM) and Spread Spectrum (SS), the results show that FCNN effectively compromises their robustness, making it a powerful yet detrimental attack method [9]. Blockchain integration provides enhanced security and traceability for digital assets, addressing concerns related to unauthorized usage [10]. Reversible watermarking methods, which allow the original image to be perfectly restored after watermark extraction, are gaining traction, especially in fields like healthcare and legal documentation.

3. Proposed Work.

In this paper, the proposed DCT-based image watermarking is proposed to generate a good imperceptibility and robustness that involves embedding and extraction process and the overall summary of the proposed method is presented in Fig. 1

3.1 *Types of Attacks*

Salt and Pepper Noise

Salt-and-pepper noise is a type of disturbance in images that manifests as random occurrences of very bright and very dark pixels scattered throughout the image. It is commonly associated with issues such as sensor malfunctions during image capture or errors in data transmission and can be represented as,

$$P'(i,j)=\begin{cases}255 & \mathrm{P}(salt)\\ 0 & P(pepper)\\ P(i,j) & P(1-P(salt)+P(pepper)\end{cases} \quad (1)$$

Where denotes the intensity at (i,j) in a clean image and is the noisy image, and denotes the probability of salt and pepper noise

Motion attack when there is relative motion between the camera and the object that distorts the captured image. This effect can either happen unintentionally, due to natural movement or environmental factors, or deliberately, to compromise the functionality of image-based systems authentication or biometric recognition as can be represented as,

$$P_{motion}(x,y)=P(x,y)*h(x,y) \quad (2)$$

Median attack is an attempt to disrupt the median filter's performance by introducing noise patterns that cause the filter to select incorrect median values, leading to distortions and by manipulating the arrangement of pixel intensities such that important details or edges are altered after filtering. The median attack can lead to effects like blurring of details, loss of structural information and misinterpretation in analysis. All these effects are potential threats to image security and can represented as,

$$P_{Noisy}(x,y)=P(x,y)+N(x,y) \quad (3)$$

3.2 *DCT*

The Discrete Cosine Transform (DCT) is a technique for transforming an input signal into the frequency domain, producing two types of coefficients: AC and DC. For example, applying DCT to an 8x8 pixel matrix yields one DC coefficient and 63 AC coefficients. The DC coefficient represents a key value related to image perception. By embedding copyright data in this coefficient, it becomes more robust against different types of manipulation, enhancing its security. The AC components comprise three types of frequencies out of which the higher frequency part is more vulnerable to attacks. The DCT transformation operation of input matrix A into output matrix B and inverse DCT can be calculated as,

$$B(x,y)=\alpha_x\alpha_y\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}A_{mn}\mathrm{Cos}\frac{\pi(2m+1)x}{2N}.\mathrm{Cos}\frac{\pi(2n+1)y}{2N} \quad (4)$$

$$A(p,q)=\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\alpha_x\alpha_y B_{xy}\text{Cos}\frac{\pi(2m+1)x}{2N}.\text{Cos}\frac{\pi(2n+1)y}{2N} \quad (5)$$

Embedding Algorithm

The algorithm used for embedding the cover image is illustrated below

(1) Input the original image (I) and watermark image (w) both are in RGB format, convert into grayscale and resize the W to match the size of I.

(2) Apply 2D-DCT to each block of I to transform it into the frequency domain by eq. 2.

(3) Select specific DCT coefficients for embedding the W and update the selected coefficient as $C'(i,j)=C(i,j)+\beta.W(i,j)$ where β is embedding strength.

(4) Apply inverse DCT as given in Eq. 5 to reconstruct the watermarked image (I_w).

Extraction Algorithm

(1) Apply 2D-DCT into 8 x 8 block of I_w.

(2) Extract watermark coefficient from selected co-efficient as,

$$W'(i,j)=\frac{C_w(i,j)-C(i,j)}{\beta}$$

(3) Reshape the W′ to match the original size of I.

4. Results and Discussion

To evaluate the performance of the DCT-based watermarking method under various types of attacks, experiments are performed on colour images as shown in Fig. 2. The size of the input image is 512 x 512 whereas the watermark has a size of 500 x 480. For measuring imperceptibility and robustness in this study the MSE parameter is used which is given in Eq. 6.

MSE

Mean Squared Error (MSE) plays a crucial role in digital image watermarking as it serves as a quantitative metric to evaluate the quality

of watermarked images and can be used to test the robustness and balance imperceptibility and strength. MSE can be calculated as,

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (P(i,j) - \hat{P}(i,j))^2 \tag{6}$$

Where M and N are the dimensions of the image, $P(i,j)$ is the pixel value of the original image at (i,j) location and $\hat{P}(i,j)$ is the pixel value of the extracted image at (i,j) location. The statistical evaluation of the different attacks on DCT-based image watermarking on MSE parameters is shown in Table 1 and visualized in Fig. 2. The original image (I) is shown in Fig. 2(a), the DCT coefficient of I is shown in Fig. 2 (b) and the watermark (W)is visualized in Fig. 2(c). The effects of attacks of salt and pepper noise, Median and motion attack after embedding W into I is shown in Fig. 2 (d,g,k). The extracted watermarks from Iw are shown in Fig. 2(e,h,l) and the original image extracted image from Iw are presented in Fig. 2(f,I,m).

In the case of salt and pepper noise the MSE of 0.0028 is obtained, whereas in the case of median attack the proposed DCT-based method has the lowest MSE value of 0.0017. In the case of motion attack, the proposed DCT-based method has obtained an MSE value higher than salt and pepper noise and lower than median attack.

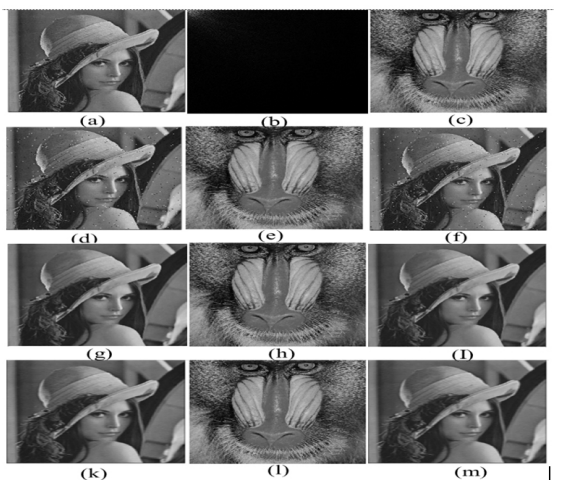

Figure 2

Visualization Effects of Noise in image watermarking (a) original image I (b) DCT of the original image (c) Watermark. Noisy image after adding salt and paper noise (d), median Attack noise (g), motion attack (k), Retrieved watermark (e,h,l), Image water removing watermark (f, I, m)

Table 1
MSE after Extracting Cover Image

Noise type/Metric	MSE
Salt and Pepper Noise	0.0028
MedianAttack	0.0017
MotionAttack	0.0021

5. Conclusion

In this paper, the efficacy of the DCT-based image watermarking method that offers a robust and effective approach to embedding copyright or authentication information within digital images is analyzed. It has been observed that DCT-based image watermarking provides a balance between imperceptibility and robustness, making it resistant to common attacks such as compression, noise, and filtering. Overall, DCT-based watermarking is a reliable technique for secure and efficient protection of digital content, maintaining both the integrity of the watermark and the quality of the image.

References

[1] F. Deeba, S. Kun, F. A. Dharejo, and H. Memon, "Digital image watermarking based on ANN and least significant bit," *Information Security Journal: A Global Perspective*, vol. 29, no. 1, pp. 30–39 (Jan. 2020), doi: 10.1080/19393555.2020.1717684.

[2] B. P. Kumari and V. P. S. Rallabandi, "Modified patchwork-based watermarking scheme for satellite imagery," *Signal Processing*, vol. 88, no. 4, pp. 891–904 (2008), doi: 10.1016/j.sigpro.2007.10.009.

[3] W. Alomoush, O. AlSheikh, A. Alsghaier, S. Al Zu'bi, and K. Alzoubi, "Digital image watermarking using discrete cosine transformation based linear modulation," *Journal of Cloud Computing*, vol. 12, no. 1, pp. 96 (2023), doi: 10.1186/s13677-023-00468-w.

[4] A. Akter, Nur-E-Tajnina, and M. A. Ullah, "Digital image watermarking based on DWT-DCT: Evaluate for a new embedding algorithm," in Proc. 2014 Int. Conf. Informatics, Electronics & Vision (ICIEV), pp. 1–6 (2014), doi: 10.1109/ICIEV.2014.6850699.

[5] M. Begum and M. S. Uddin, "Implementation of secured and robust DFT-based image watermark through hybridization with decomposition algorithm," *SN Computer Science*, vol. 2, no. 3, pp. 221 (2021), doi: 10.1007/s42979-021-00608-6.

[6] A. Rai and H. V. Singh, "SVM based robust watermarking for enhanced medical image security," *Multimedia Tools and Applications,* vol. 76, no. 18, pp. 18605–18618 (2017), doi: 10.1007/s11042-016-4215-3.

[7] R. Mehta, V. P. Vishwakarma, and N. Rajpal, "Lagrangian support vector regression based image watermarking in wavelet domain," in Proc. 2015 2nd Int. Conf. Signal Processing and Integrated Networks (SPIN), pp. 854–859 (2015), doi: 10.1109/SPIN.2015.7095436.

[8] H. K. Singh and A. K. Singh, "Digital image watermarking using deep learning," *Multimedia Tools and Applications,* vol. 83, no. 1, pp. 2979–2994 (2024), doi: 10.1007/s11042-023-15750-x.

[9] M. W. Hatoum, J.-F. Couchot, R. Couturier, and R. Darazi, "Using deep learning for image watermarking attack," *Signal Processing: Image Communication*, vol. 90, pp. 116019 (2021), doi: 10.1016/j.image.2020.116019.

[10] B. Wang, S. Jiawei, W. Wang, and P. Zhao, "Image copyright protection based on blockchain and zero-watermark," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 4, pp. 2188–2199 (2022), doi: 10.1109/TNSE.2022.3157867.

Received March, 2025