Acadlore Transactions on AI and Machine Learning
https://www.acadlore.com/journals/ATAIML

DWT-Based Digital Watermarking for Various Attacks

Nail Alaoui*

Laboratoire de Recherche Modélisation, Simulation et Optimisation des Systèmes Complexes Réels, Université Ziane Achour de Djelfa, 17000 Djelfa, Algeria

* Correspondence: Nail Alaoui (n.alaoui@univ-djelfa.dz)

Received: 09-07-2023 **Revised:** 10-28-2023 **Accepted:** 11-05-2023

Citation: N. Alaoui, "DWT-based digital watermarking for various attacks," *Acadlore Trans. Mach. Learn.*, vol. 2, no. 4, pp. 226–234, 2023. https://doi.org/10.56578/ataiml020405.

Abstract: In the domain of intellectual property protection, the embedding of digital watermarks has emerged as a pivotal technique for the assertion of copyright, the conveyance of confidential messages, and the endorsement of authenticity within digital media. This research delineates the implementation of a non-blind watermarking algorithm, utilizing alpha blending facilitated by discrete wavelet transform (DWT) to embed watermarks into genuine images. Thereafter, an extraction process, constituting the inverse of embedding, retrieves these watermarks. The robustness of the embedded watermark against prevalent manipulative attacks, specifically median filter, salt and pepper (SAP) noise, Gaussian noise, speckle noise, and rotation, is rigorously evaluated. The performance of the DWT-based watermarking is quantified using the peak signal-to-noise ratio (PSNR), an objective metric reflecting fidelity. It is ascertained that the watermark remains tenaciously intact under such adversarial conditions, underscoring the proposed method's suitability for applications in digital image security and copyright verification.

Keywords: Digital watermarking; Discrete wavelet transform; Robustness; Alpha blending; Peak signal-to-noise ratio

1 Introduction

The advent of digital technology has facilitated the storage and manipulation of media content, simultaneously escalating concerns regarding security. In the milieu of digital media, the integration of digital watermarks has been recognized as an efficacious strategy to fortify copyright protection [1]. The practice of digital watermarking, characterized by the insertion of ownership data into multimedia files, enables the subsequent retrieval of this information to ascertain the legitimate proprietor of the content.

Watermarking fulfills multiple objectives: It is instrumental in verifying digital image integrity, safeguarding concealed data, mitigating unauthorized replication and distribution over the internet, and affirming ownership. Within the realm of watermarking techniques, a dichotomy is observed: the spatial domain and the frequency domain. The former relies on the perceptual analysis of the image, subtly modifying pixels within select image subsets, while the latter targets specific frequency components for alteration based on pre-established criteria [2]. Frequency domain methods, known for their robust and discreet watermarking effects, are predominantly preferred over their spatial domain counterparts.

The taxonomy of watermarking algorithms bifurcates into blind and non-blind techniques, distinguished by their extraction requisites. Non-blind watermarking mandates the presence of the original image for watermark extraction, unlike its blind counterpart. The present research adopts a non-blind approach, necessitating the original image for the extraction process. It is dedicated to devising a real-time image authentication system anchored in the DWT domain watermarking [3]. The robust watermarking technique being proposed utilizes the DWT, leveraging its capability for multiresolution analysis and spatial localization, aptly aligning with the overarching goals of digital image security and copyright protection.

The significance of this study is encapsulated in its contribution to the field of digital image authentication and copyright protection. By amalgamating DWT with alpha blending, an equilibrium between image quality and watermark robustness is sought, advancing the corpus of existing watermarking methodologies.

https://doi.org/10.56578/ataiml020405

2 Related Work

The domain of digital image watermarking is rich with methodologies that bifurcate into two primary realms: the spatial and frequency domains [4]. Within the spatial domain, watermarking is accomplished by subtle alterations to pixel values [5], whilst frequency domain watermarking necessitates the transformation of both host and watermark images into the frequency spectrum prior to modification. For color images, a probabilistic, block-based method is applied in the spatial domain [6].

Recent advancements have seen binary watermarking techniques gaining traction, leveraging sequence numbers derived from secret keys and Gray code [7]. Least significant bit (LSB) watermarking is prevalently utilized, embedding watermark information into the more complex areas of the host image. This method navigates the balance between high capacity and low perceptibility [8].

In the non-blind watermarking paradigm, the DWT in tandem with the discrete cosine transform (DCT) is employed, characterized by its imperceptibility, and obviating the need for the original image during extraction [9]. DCT-based watermarking has been distinguished by its practice of embedding the watermark centrally within the host image [10].

The introduction of PPLU decomposition to the watermarking arena presents a fortified method, deploying mathematical transformations to safeguard the watermark [11]. Block-based watermarking approaches have proved beneficial in scenarios where additional space is necessitated for multiple watermarks, thereby enhancing capacity [12].

For bolstering robustness, the DWT-singular value decomposition (DWT-SVD) technique has been embraced. This sophisticated approach provides enhanced security against various image processing attacks [13]. It should be noted that while each method exhibits distinctive merits, the current research has elected to utilize the combination of DWT and alpha blending. This fusion is carefully selected to fulfill the imperceptibility criterion essential for practical digital image watermarking applications.

3 Digital Watermarking

Within the field of digital watermarking, two pivotal phases are delineated: the embedding and the extraction of the watermark. As illustrated in Figure 1, the cover image undergoes a modification process during embedding, resulting in the generation of a watermarked image. The extraction phase, depicted in Figure 2, involves the retrieval of the watermark from the watermarked image. The resilience of the watermarked image is assessed by subjecting it to an array of attacks, and the effectiveness of the watermarking is subsequently evaluated through a comparison of the extracted watermark with the original watermark and the cover image.

The hallmarks of an efficacious watermarking technique are imperceptibility, robustness, capacity, and security:

Imperceptibility is ensured through the application of alpha blending, which facilitates the seamless integration of the watermark into the host image, maintaining the visibility at a controlled level.

Robustness is achieved by employing DWT in the frequency domain, which provides a multiresolution analysis and fortifies the watermark against common attacks.

Capacity is optimized by striking a delicate balance between the watermark's volume and the perceptual transparency through selective embedding techniques.

Security is enhanced through the use of non-blind watermarking and DWT-based embedding, safeguarding the watermark's integrity.

In the methodology presented, the implantation involves the modification of DWT coefficients to embed the watermark, and the extraction process entails the precise retrieval of these coefficients. This bidirectional approach is pivotal to the establishment of a robust and effective watermarking strategy.

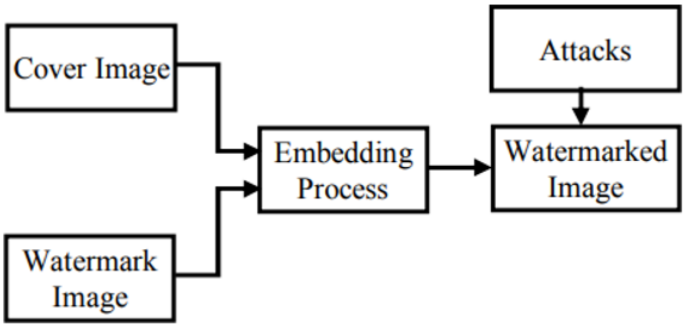


Figure 1. Embedding process

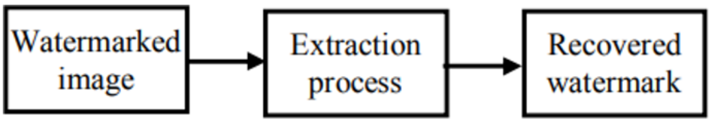


Figure 2. Extraction process

Figure 1 and Figure 2 serve as block diagrams that elucidate the embedding and extraction processes, respectively, within the watermarking workflow.

4 DWT

DWT is recognized as a hierarchical decomposition technique in the realm of mathematical transformations. This process is characterized by the deconstruction of a signal into an assembly of wavelets, brief oscillatory waveforms of varied frequencies. It is the attributes of these wavelets that facilitate the division of the original signal into a series of transform coefficients, each encapsulating distinct positional data. Through the application of an inverse wavelet transformation, these coefficients permit the full reconstruction of the initial signal.

The selection of DWT for watermarking methodologies is substantiated by its provision of multiresolution analysis and its capacity for energy compaction and spatial localization. Additionally, the orthogonality and prevalent use of DWT offer an equilibrium between robustness and transparency that is both practical and efficient for the incorporation of watermarks into images.

It is the distribution of information across disparate frequency subbands by DWT that underpins the robustness of watermarks, ascertaining their detectability despite potential alterations within specific subbands resulting from image processing or attacks.

A breakdown of an image via DWT yields three predominant components, constituting one approximation subband and two detailed subbands or subbands, designated as LL, LH, HL, and HH bands, as demonstrated in Figure 3. The LL band comprises low frequencies in both the horizontal and vertical planes, while the HH band captures high frequencies across the same axes. Conversely, the HL band encompasses high horizontal frequencies coupled with low vertical frequencies, and the LH band, high vertical frequencies alongside low horizontal frequencies. The low-frequency portion harbors rudimentary information regarding the signal, whereas the high-frequency portion is replete with intricate details pertaining to edge components.

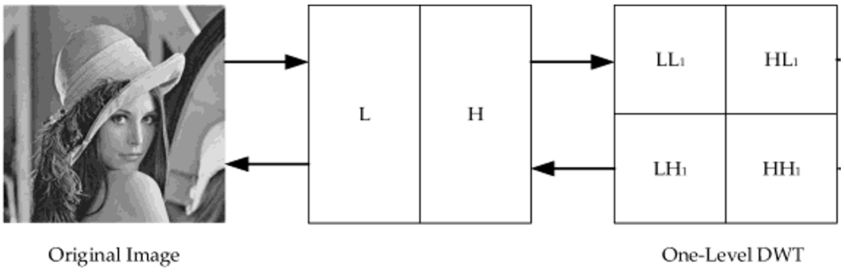


Figure 3. One-level discrete wavelet decomposition

In the context of image energy proportionality and image approximation representation, the LL band is deemed significantly important. Conversely, the high-frequency detail bands, namely, LH, HL, and HH, are less discernible to the human eye. This characteristic renders them as favorable locales for watermark embedding, as alterations within these regions are less apparent. Embedding in these less conspicuous areas is strategic, bolstering the watermark's tenacity whilst minimally impairing image quality.

In the decomposition process at each level, DWT is executed bidirectionally, initially in the vertical orientation, followed by the horizontal. The prime decomposition stage produces four subbands: LL_1, LH_1, HL_1, HH_1. Each successive decomposition level utilizes the preceding LL subband as the input, which is further subdivided into four new multi-resolution subbands. This generates a progressively coarser set of wavelet coefficients for subsequent levels, the number of which is contingent upon the specific application.

Owing to its stellar performance in localizing perturbations within both spatial and frequency dimensions, DWT has ascended in popularity for its efficacy in camouflaging disturbances within discrete image sectors. This technique, not necessitating the original image for watermark extraction, finds versatility in a spectrum of signal processing applications, encompassing denoising, as well as audio and video compression. The methodology of digital image watermarking is primarily comprised of two phases: the embedding of the watermark into the data and the subsequent extraction thereof.

The incorporation of DWT in the watermarking methodology is instrumental due to its division of the image into distinct subbands, each with unique characteristics that facilitate the embedding of the watermark. The utility of DWT lies in its capacity for multiresolution analysis, precise localization, concentration of energy within the signal, and its reversible nature, ensuring that the watermark can be tailored to the image content and robust against various attacks while maintaining the integrity of the original image quality.

a. Watermark embedding

Upon application of a two-dimensional DWT to the cover image, four sub-bands are yielded: the low-frequency approximation, high-frequency diagonal, low-frequency horizontal, and low-frequency vertical sub-bands. The watermark is also subjected to a two-dimensional DWT using the Haar wavelet. Alpha blending is utilized for the embedding process, wherein the watermark is applied by scaling the separate components of the watermark and cover image by predetermined factors before their amalgamation. The embedding procedure stipulates that the dimensions of the cover image exceed those of the watermark, albeit requiring the frame sizes to be congruent.

For visible watermarking applications, the watermark is integrated into the low-frequency approximation component of the cover image.

b. Alpha blending technique for watermark embedding

The technique of alpha blending for embedding the watermark in the image employs the following equation:

$$WMI = k^*(LL_1) + q^*(WM_1) \tag{1}$$

where, *WMI* is defined as the watermarked image, LL_1 as the low-frequency approximation of the original image, WM_1 as the watermark, with k and q representing the scaling factors for the original and watermark images, respectively.

The secure watermarked image is then reconstructed through the execution of an inverse DWT on the watermarked image coefficients.

Figure 4 illustrates the methodology employed for the watermark embedding.

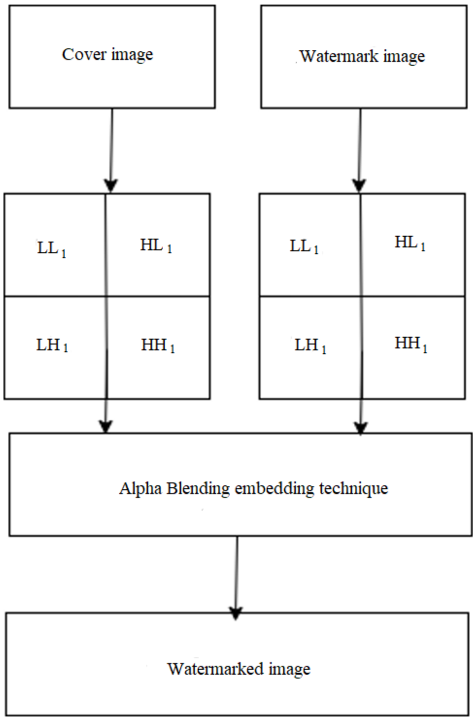


Figure 4. Watermark embedding technique

c. Watermark extraction

The extraction of the watermark is executed by inverting the stages of the embedding process. Both the cover and watermarked images are subjected to the initial DWT. The watermark is then isolated from the watermarked image by employing an alpha compositing technique.

d. Alpha blending technique for watermark extraction

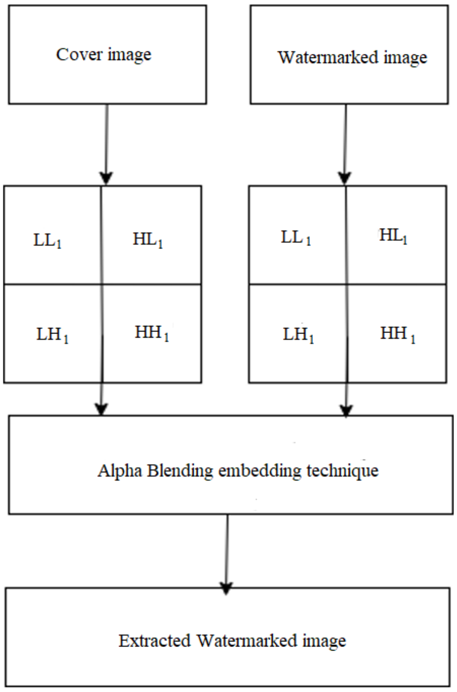
The extraction of the watermark utilizes the alpha blending technique according to the equation:

$$RW = (WMI - k^{*}LL_1) \tag{2}$$

where, *RW* denotes the recovered watermark, *WMI* is the watermarked image, LL_1 represents the low-frequency approximation of the original image.

The final watermark extraction image is derived by performing an inverse DWT on the coefficients of the watermark image.

Figure 5 delineates the technique employed for the watermark extraction.



Cover image
Watermarked image
LL1
HL1
LH1
HH1
LL1
HL1
LH1
HH1
Alpha Blending embedding technique
Extracted Watermarked image

Figure 5. Watermark extraction technique

5 Results

The experimental section utilized MATLAB 2018a to simulate the proposed digital image watermarking scheme based on DWT. Standard images of dimensions 512×512 pixels served as the testbed to ascertain the algorithm's resilience. The images, once embedded with the watermark using the established technique, were exposed to a battery of manipulations to evaluate robustness.

To evaluate the robustness of the watermarked image, it was subjected to a series of perturbations. These perturbations included the addition of Gaussian noise, speckle noise, and SAP noise, as well as the application of rotation attack and median filter.

To this end, the watermarked images underwent deliberate perturbations, which encompassed a spectrum of noise types and processing attacks, selected for their prevalence in real-world scenarios where image integrity is often compromised. These included:

Gaussian noise: The addition of Gaussian noise tested the watermark's durability against distortions commonly encountered during transmission and storage.

Speckle noise: The introduction of speckle noise provided an assessment of the watermark's tenacity against granular interference, characteristic of this noise type.

SAP noise: By introducing this impulse noise, the watermark's resilience to random pixel corruption was gauged.

Median filter: The application of median filter, a staple in image denoising, challenged the watermark's persistence through common image enhancement processes.

Rotation: The rotation of watermarked images at various angles assessed the watermark's stability in the face of geometric distortions.

The efficacy of the watermarking process was quantified using *PSNR*, a measure of the watermark's perceptibility post-recovery. Objective quality metrics were employed to evaluate the fidelity of the recovered image following the application of filters. The mean square error (*MSE*) and the *PSNR* values were computed according to the established equations, comparing the original (*O*) and the recovered image (*R*), with *N* and *M* representing the image's dimensions.

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \tag{3}$$

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[O(i,j) - R(i,j)]^2 \tag{4}$$

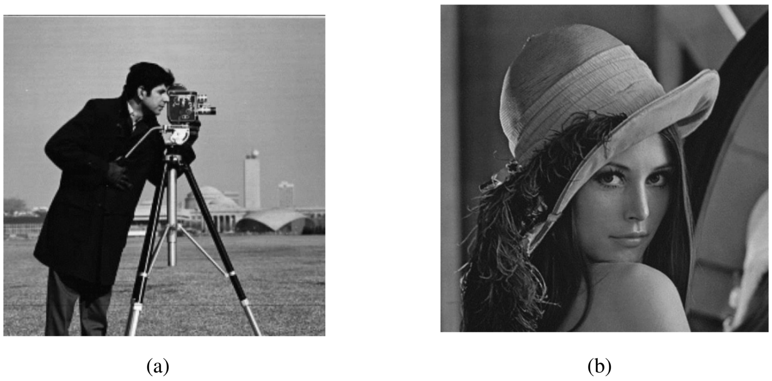


Figure 6. (a) Cover image; (b) Watermark image

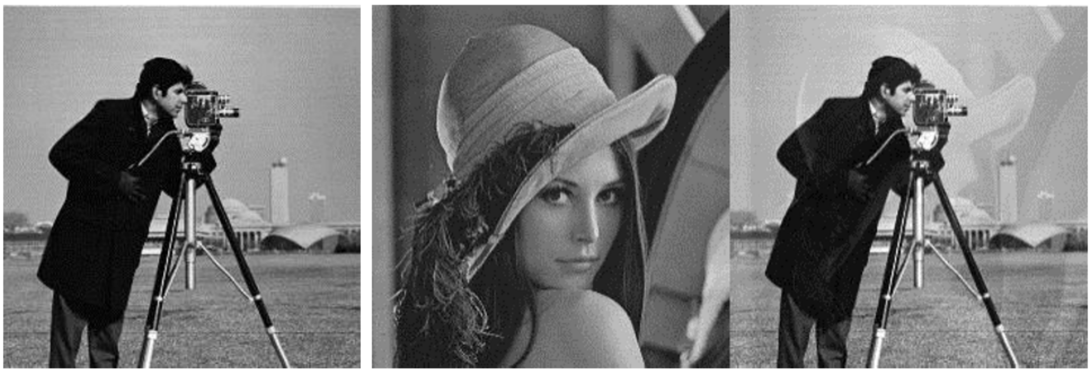


Figure 7. One-level DWT watermark embedding procedure

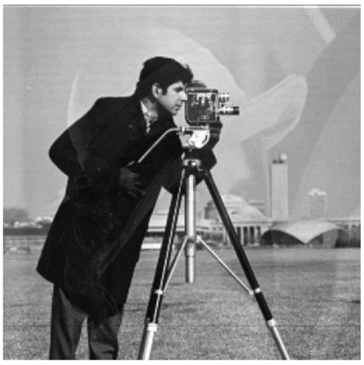


Figure 8. Watermarked image

Subgraphs (a) and (b) of Figure 6 depict the original cover image and watermark image respectively. The embedding process, illustrated in Figure 7, demonstrates the watermark insertion at one-level DWT. The watermarked image, as evidenced in Figure 8, highlights the outcomes of varying the parameter k from 0.1 to 0.9 while maintaining the constant q at 0.1. An optimal balance is achieved at k equal to 0.5, whereas values below 0.4 resulted in the watermark dominating the visual hierarchy of the cover image, leading to an undesirable prominence.

In the undertaking of this study, the experimental procedures were designed to adhere to the rigorous standards expected by leading academic publications. The passive voice has been predominantly employed to align with the formal style of such journals, and terminology has been standardized to ensure consistency throughout the manuscript.

Extraction of the watermark was visualised in Figure 9, illustrating the one-level DWT watermark extraction process. Subsequently, Figure 10 delineates the variations in the appearance of the extracted watermark as parameter k was modified from 0.1 to 0.9. It was observed that a decrement in k to values below 0.2 precipitated a darkening of the watermark, culminating in its complete absence from the visual domain.

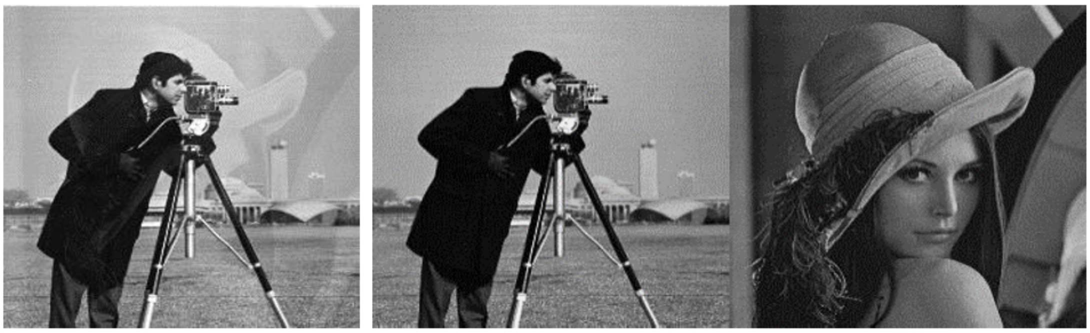


Figure 9. One-level DWT watermark extraction process

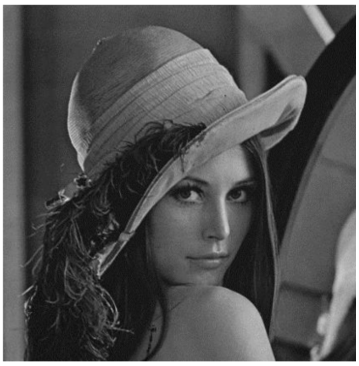


Figure 10. Extracted watermarked image

Table 1. Watermarked and isolated watermark images after attacks

No.	Attacks	PSNR (dB)
1	SAP noise	35.0125
2	Median filter	33.4795
3	Rotation	11.0025
4	Speckle noise	33.2976
5	Gaussian noise	25.4510

Table 1 encapsulates the performance of both watermarked and the isolated watermark images under the duress of several image processing attacks, with the corresponding PSNR values detailed. The resilience of the watermarking approach is quantitatively substantiated by the PSNR values: 35.0125 dB for SAP noise, 33.4795 dB for median filter, 33.2976 dB for speckle noise, 25.4510 dB for Gaussian noise, and 11.0025 dB for rotation attack. These values suggest that the watermark remains discernible, notwithstanding the application of various distortions, thereby evidencing its durability.

Further, Table 2 presents a granular breakdown of the watermarked image's robustness across various densities of SAP noise, Gaussian noise, and speckle noise, with a direct correlation observed between increased noise density and diminished PSNR values. Notably, as noise density lessens, the robustness of the watermarked image correspondingly strengthens.

Table 2. Various noise values in watermarked image with corresponding PSNR

Attacks	Density	PSNR
SAP noise	0.1	15.3058
	0.01	24.9824
	0.001	35.0125
Gaussian noise	0.1	16.5201
	0.01	19.9850
	0.001	25.4510
Speckle noise	0.1	15.7852
	0.01	21.7489
	0.001	33.2976

These PSNR metrics are pivotal, providing objective and quantifiable evidence of the watermarking method's robustness against diverse image manipulations. Such empirical data are indispensable for ascertaining the method's efficacy in practical scenarios pertaining to image authentication and copyright protection.

In synthesis, the battery of tests conducted demonstrates the proposed watermarking technique's capacity to withstand an array of image processing attacks. Supported by both quantitative PSNR metrics and qualitative visual analysis, the technique evidences a high degree of resilience, thus positioning it as a viable contender for real-world application in safeguarding digital image integrity. Albeit promising, the empirical outcomes suggest a necessity for further validation in operational environments to fully endorse the utility of this watermarking approach in image security applications.

6 Conclusions

In the current investigation, a DWT-based alpha blending watermarking algorithm has been developed and assessed. It has been demonstrated that through the application of this method, a watermark is conspicuously embedded within the cover image, with its extraction necessitating the original image. The fidelity of the embedded watermark as well as the watermarked image is governed by the scaling parameters k and q.

A series of robustness tests were conducted, wherein the watermarked images were subjected to various manipulative attacks, including Gaussian noise, speckle noise, SAP noise, rotation, and median filter. The findings reveal a notable resilience of the watermarking method against these common image processing challenges.

Future work will pivot towards the enhancement of this algorithm, exploring the integration of diverse transformation techniques. These will encompass geometric transformations, frequency domain transformations, and multi-resolution watermarking strategies. The intent is to further fortify the watermark against a spectrum of manipulations and infringements, ensuring the watermark's integrity and detectability even when subjected to strenuous conditions.

The efficacy of these advanced techniques is anticipated to vary, contingent upon the intrinsic properties of the images and the nature of the attacks. This variability underscores the necessity for an expansion of empirical research, with a focus on extensive experimentation and rigorous data analysis.

Data Availability

The data used to support the research findings are available from the corresponding author.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] A. Mohanarathinam, S. Kamalraj, G. K. D. Prasanna Venkatesan, R. V. Ravi, and C. S. Manikandababu, "Digital watermarking techniques for image security: A review," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, pp. 3221–3229, 2020. https://doi.org/10.1007/s12652-019-01500-1

[2] M. Begum and M. S. Uddin, "Digital image watermarking techniques: A review," *Information*, vol. 11, no. 2, p. 110, 2020. https://doi.org/10.3390/info11020110

[3] N. Fatès and F. A. P. Petitcolas, “Watermarking scheme evaluation tool,” in *Proceedings International Symposium on Multimedia Software Engineering, Taipei, Taiwan*, 2000, pp. 328–331. https://doi.org/10.1109/MMSE.2000.897231

[4] J. Licks and R. Jorden, “Geometric attacks on image watermarking system,” *IEEE Multimedia*, vol. 12, no. 3, pp. 68–78, 2005. https://doi.org/10.1109/MMUL.2005.46

[5] R. Tay and J. P. Havlicek, “Image watermarking using wavelets,” in *2002 45th Midwest Symposium on Circuits and Systems, MWSCAS-2002, Tulsa, OK, USA*, 2002. https://doi.org/10.1109/MWSCAS.2002.1187021

[6] V. M. Potdar, S. Han, and E. Chang, “A survey of digital image watermarking techniques,” in *INDIN ’05. 2005 3rd IEEE International Conference on Industrial Informatics, Perth, WA, Australia*, 2005, pp. 709–716. https://doi.org/10.1109/INDIN.2005.1560462

[7] S. P. Maity and M. K. Kundu, “Robust and blind spatial watermarking in digital image,” in *ICVGIP 2002, Proceedings of the Third Indian Conference on Computer Vision, Graphics & Image Processing, Ahmadabad, India*, 2022.

[8] I. Nasir, Y. Weng, and J. M. Jiang, “A new robust watermarking scheme for colour image in spatial domain,” in *2007 Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, Shanghai, China*, 2007. https://doi.org/10.1109/SITIS.2007.67

[9] R. K. Arya, S. Singh, and R. Saharan, “Secure non-blind block based digital image watermarking technique using DWT and DCT,” in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India*, 2015, pp. 2042–2048. https://doi.org/10.1109/ICACCI.2015.7275917

[10] R. Verma and A. Tiwari, “Copyright protection for watermark image using LSB algorithm in colored image,” *Adv. Electron. Electr. Eng.*, vol. 4, no. 5, pp. 499–506, 2014.

[11] N. Muhammad and N. Bibi, “Digital image watermarking using partial pivoting lower and upper triangular decomposition into the wavelet domain,” *IET Image Process.*, vol. 9, no. 9, pp. 795–803, 2015. https://doi.org/10.1049/iet-ipr.2014.0395

[12] M. Khan, A. Kushwaha, and T. Verma, “A new digital image watermarking algorithm based on image interlacing, DWT, DCT,” in *2015 International Conference on Industrial Instrumentation and Control (ICIC), Pune, India*, 2015, pp. 885–890. https://doi.org/10.1109/IIC.2015.7150868

[13] P. Priyanka and S. Maheshkar, “An efficient DCT based image watermarking using RGB colour space,” in *2015 IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS), Kolkata, India*, 2015. https://doi.org/10.1109/ReTIS.2015.7232881