JOURNAL OF SOFT COMPUTING AND DATA MINING
e-ISSN: 2716-621X

JSCDM

Vol. 6 No. 1 (2025) 435-447
https://publisher.uthm.edu.my/ojs/index.php/jscdm

# Employing Hybrid Watermarking to Improve Email Security Against Cyber Attacks

**Ibrahim M. Ahmed[1]*, Awos Kh. Ali[2], Mahmood S. Mahmood[3]**

[1] *Department of Cybersecurity, College of Computer Science and Mathematics, University of Mosul, Mosul, IRAQ*
[2] *Department of Computer Science, College of Education for Pure Science, University of Mosul, Mosul, IRAQ*
[3] *Department of Computer Science, College of Science, University of Mosul, Mosul, IRAQ*

*Corresponding Author: ibrahim_alhlima@uomosul.edu.iq
DOI: https://doi.org/10.30880/jscdm.2025.06.01.029

**Article Info**

Received: 26 December 2024
Accepted: 18 April 2025
Available online: 30 June 2025

**Keywords**

Cyberattacks, email security, watermarks, invisible watermark, visible watermark, networks

**Abstract**

Email security is increasingly essential due to the growing number of IT security threats. This paper aims to assess the risks associated with emailing as a communication medium, including message interception, message loss, message content tampering, and intellectual property theft. These threats can be mitigated by employing Watermarking (WM) techniques, which have become a highly effective data protection tool. Information placed in digital content can be hidden, which can be used for document authentication and should not be exposed to unauthorized users. Hybrid watermarking is an advanced approach that can be exploited to enhance the security of an email, where both visible and invisible marks are strategically placed to protect the integrity, authenticity, and confidentiality of email content. This integration provides a robust security model that can prevent unauthorized access and identify any unauthorized use or tampering with digital content. In this paper, invisible watermarking is employed to enhance email security, complementing visible watermarks. This approach aims to strengthen email protection by increasing the likelihood of easily identifying attack attempts to compromise content or unauthorized retrieval of secure data from received emails. The performance of the selected approaches has been assessed by employing some forms of attacks, including scaling, reformatting, denoising, and noise reduction. Several metrics have been utilized to validate the quality of tested files: the cover image file's PSNR for Peak Signal to Noise Ratio, the invisible watermark file's BER and NC for Normalized Correlation.

## 1. Introduction

Nowadays, many applications have emerged, and the majority of them are dominated by communication technology; however, email remains one of the most effective legal methods of passing information in business and other professional activities. Nevertheless, there are many concerns regarding the security of email content, as hackers are widely utilizing advanced, emerging cyber threats. Regarding the protection of email communication, concerns exist about hacking, information leakage, message content tampering, and piracy, regardless of the various types of communication mediums [1].

Thus, in relation to these threats, distinct forms of watermarking methods have turned into an effective tool for data protection, leveraging their extracted features [2]. Computerized content may have visible or hidden text, images, and codes for identification, ownership, and tracking purposes. Overall, the two most familiar forms of watermarking techniques are visible and invisible watermarking. Any content with visible watermarks in the form of text or logos may be easily recognized upstream [3]. As a sign of ownership or legitimacy, they are conspicuous as far as the eye can see. On the other hand, it is inserted into the information and cannot be seen by the patient; this can only be erased using particular algorithms or even other software. However, the Watermarking technique could face vulnerabilities to attacks such as compression, cropping, scaling, noise, or format conversion, as well as difficulties in maintaining watermark integrity after signal processing operations.

Image transmission has become popular in daily life over the Internet. Most Internet applications and social media networks have image transmission functions; therefore, these applications and authorities should provide obligations to keep these images secure from malicious parties. Realizing that transmitted images, which often have a higher resolution, require a large amount of storage space. The issue, though, is that there are indeed opportunities for cost and time efficiencies within the redundancy of the digital media transmission. At the same time, it permits concealing information by altering the raw information of images, thereby eliminating the influence of the Least Significant Bit (LSB) [4]. Especially in military affairs, commercial, financial, and secret negotiations, agreements, and contracts between negotiating parties, the transfer of a large number of messages through the Internet in digital media is of great importance. On the other side, there is also the potential of fake media that is very difficult to detect with the naked eye or even a simple detection mechanism. Thus, making accurate conclusions about the authorship of data in communications conducted on the Internet gains particular importance in cases of conflict of interest. For the protection of content security in digital media, several solutions have been proposed for various uses, including copyright protection, fingerprinting, and authentication [5]. Regarding copyright protection and authentication, cryptographic, steganographic, and watermarking methods have been established as the most effective and practical methods for protecting the rights of digital multimedia [3].

Additional component indications for watermarking include issues such as strength and invisibility. The term imperceptibility defines the physical appearance of images. If the exact perception of the original and embedded image is matched, then it is quite subtle. It has been usual to describe watermark pictures as robust by pointing out how close they are to the original pictures [6, 7]. This paper demonstrates that implementing a single method of imperceptible watermarking that is robust and has maximum embedding capacity simultaneously is inefficient in terms of security. The watermark should be able to resist various types of malicious or unintended image processing operations, which constitutes the meaning of robustness [8]. It is quite possible to quantify the models' ability to hold up in general and the precision of the extraction following different assaults. Its removal often employs a plethora of methods that fall into two broad categories: geometric attack and signal processing attack. There have also been a number of newly publicized assaults[6]. The bit error rate (BER) and the normalized cross-correlation (NCC) are two commonly used indices to evaluate resilience [9, 19].

Each of the earlier-mentioned watermarking methods has its advantages and weaknesses [10]. Static, straight, and easily noticeable watermarks are effective to prevent unauthorized usage and copying by customers, yet they are also easy to be wiped out or rewritten by attackers. Furthermore, invisible watermarks are not always protected from attacks that can manipulate or alter the material without the observer's knowledge. The area of digital security has recently shown interest in a compound type of watermarking that combines visible and invisible watermarks, as each has its shortcomings [11].

Private documents are usually mailed; these emails have to be protected against hacks and various types of attacks to ensure that no data is lost or manipulated. While using email content, it is possible to encrypt or sign it; however, this does not eliminate cases of tampering or leakage [12]. Visible and invisible watermarks can be beneficial for protecting emails. The recognition speed of the visible watermark is greatly enhanced, and covert tracking information is secure even if the visible watermark is distorted. Hybrid watermarking not only prevents unauthorized use but also ensures content integrity and authenticity during distribution. Thus, the hybrid watermarking method has two advantages, specifically for email security [13]. The overall ideas behind both visible and invisible watermarks do not depend on one another; it is more challenging for attackers to manipulate both at the same time, thus improving the combination's resilience. One of the clear and immediate signs of ownership or copyright, the visible watermark hinders illicit use. One approach is to employ discrete tracking and verification, which can be facilitated with the aid of invisible watermarks. The purpose of employing Hybrid Watermarking (HWM) in enhancing email security is to establish a robust security framework that addresses the limitations of traditional watermarking techniques [13].

The main aim of this paper is to utilize both visible and invisible watermarking to fulfill the following objectives:

- Improve robustness of watermarks against various types of attacks, including cropping,
- compression, and others, and improving the ability to modify watermarks easily.
- Improve data Integrity to ensure that the confidentiality of information transferred by email is preserved, so that if someone changes it, the change can be identified.
- Preserve the proprietary content in that it should be traceable and securely distinguishable from its digital content.
- Deter unauthorized distribution.

## 2. Related Work

Many data protection technologies have been proposed to enhance security; one notable technology is watermarking. This technique has two broad categories, namely the visible and the invisible watermarking. The former is for copyright issues, and the latter is for authentication purposes, which is the concern of this paper. The three basic properties of invisible watermarking are transparency, capacity, and robustness[14]. First, when adding watermark data to a cover file, the quality of the cover file cannot easily be judged by human senses, which means that there can be a number of instances where more watermark data may be injected without really affecting the special feature, the capacity. Lastly, being robust is crucial in this case because a possibility is demonstrated that the techniques used can make the watermark cover file scalable, reformatted, noisy, and reconstructed at the receiving end, yet still remain readable [18].

Mohammed Hassan Abd and Osamah Waleed Allawi [15] presented an example of an invisible watermarking technique using the Least Significant Bit (LSB) method, which operates in the spatial domain. In contrast, other approaches, such as those based on the Discrete Wavelet Transform (DWT), work within the frequency domain. Transform or frequency domain techniques have small capacity and high robustness, spatial domain approaches have huge capacity and fragile robustness against malicious assaults. Potentially valuable for the imperceptibility measurements, this may be determined with the PSNR (Peak Signal to Noise Ratio) [9].

Lita Lidyawati et al. [9] developed a sorting watermarking application that heavily relies on the PSNR value, as it depends on the degree of distortion and the quality required by the application. The value of PSNR will recommend the type of transform. Invisible watermarking is uncovering hidden data within a medium that nobody can detect. The primary concern of undetectable watermarking is to enhance the system's indistinguishability and robustness. Researchers have proposed numerous security methods to enhance the watermarking system. However, most of these methods require strong anti-piracy legislation to build safeguards to be secure[9].

Ali Fatahbeygi and Fardin Akhlaghian Tab [10] developed a highly robust image watermarking technique, demonstrating strong resistance against various image processing attacks. This process involves intentionally hiding data within an image. Additionally, cryptography is another means of coding and sending data, which prevents unauthorized parties from accessing it. Invisible watermarking refers to both spatial domain watermarking and frequency domain watermarking. Frequency domain watermarking is more resilient when compared with spatial domain watermarking[10].

Abdulhakeem O. Mohammed et al. [12 ] used the Discrete Cosine Transform (DCT), which assists in further isolation of coefficients in a low-frequency subband, and they are more correlated with one another. After DCT, a large amount of energy is concentrated throughout the entire picture, as only a few out of a large number of coefficients carry most of the low-frequency subband energy. This would have altered the image more if these coefficients had not been adjusted consistently, so it should have ensured that they were not adjusted. Watermarking robustness and simplicity can be made consistent, as high coefficients aren't visually delicate, and the high-frequency sections of the low-frequency subband are the ideal place to embed watermarks[12].

Zhen Dai et al. [11] proposed a novel framework that combines reversible watermarking with zero watermarking, aiming to enhance both security and reversibility in digital media protection. In the first part, an ownership share is constructed using watermark data and characteristics derived from the closest neighbor greyscale residual (NNGR). In the second part, we have reversibly embedded the previously created ownership share with Slantlet Transform, Singular Value Decomposition, and Quantization Index Modulation (SLT-SVD-QIM). Our proposed scheme is compared to other medical image watermarking schemes in terms of both watermark imperceptibility, watermark distinguishability, and watermark robustness, as well as a continuous verification function provided without resorting to dispute resolution or third-party storage [10].

Radha Kumari et al.[16] that utilizes the DWT-SVD approach to detect and counter various watermark attacks on the digital watermark technique. The purpose of this method is to find a connection between invisibility and resilience using scaling variables. This study aims to embed and extract multiple attacks on watermarked images without compromising the host or the watermarked picture [16]. Most of the recent literature has not been tested on various types of cyber attacks. In this paper, we exploit both visible and invisible watermarks in a manner that

operates independently to mitigate, manipulate, and enhance the combination's resilience. Table 1 summarizes the related work.

**Table1** *Summary of the related work*

| Author(s) | Technique(s) | Domain | Main Focus | Strengths | Weak Points |
|---|---|---|---|---|---|
| Mohammed H. Abd & Osamah W. Allawi [15] | LSB, DWT | Spatial & Frequency | Compares spatial and frequency domain watermarking methods | Spatial domain offers high capacity; frequency domain has better robustness | Spatial methods are easy to break; frequency methods hold less data |
| Lita Lidyawati et al. [9] | PSNR-based watermarking | Not clearly specified | Emphasizes the role of PSNR in choosing the right watermarking technique | Helps match method with image quality needs | Depends heavily on legal protections to stay secure |
| Ali Fatahbeygi & Fardin A. Tab [10] | Cryptography + watermarking | Spatial & Frequency | Focuses on making watermarking more secure and harder to tamper with | Strong security and robustness using encryption and frequency methods | Spatial methods still less resilient; details on limitations not clear |
| Abdulhakeem O. Mohammed et al. [12] | Discrete Cosine Transform (DCT) | Frequency | Embeds watermark in low-frequency parts of the image to ensure stability | Maintains image quality while staying robust | Needs careful tuning to avoid noticeable changes in the image |
| Zhen Dai et al. [11] | Reversible + Zero Watermarking (SLT-SVD-QIM) | Frequency | Designed for medical images with focus on ownership proof and imperceptibility | Highly robust, supports continuous verification without third parties | Technique is quite complex and may require more processing power |
| Radha Kumari et al. [16] | DWT-SVD | Frequency | Balances between invisibility and strength against various attacks | Handles multiple types of attacks without affecting image quality | Finding the right scaling factors can be tricky |

## 3. Proposed Robust HWM System

The aim of the proposed method is to insert watermarks inside or as an attachment to an email message or email attachment (for example, PDF files, pictures, or documents) by applying the visible watermarking to the original invisible watermarking in the email. The proposed Framework improves the security of email against cyber attacks by making any means of change or theft attempts on emails more recognizable and detectable. The proposed methodology operates with multiple steps in the sending side of email as depicted in Fig 1.

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 435-447

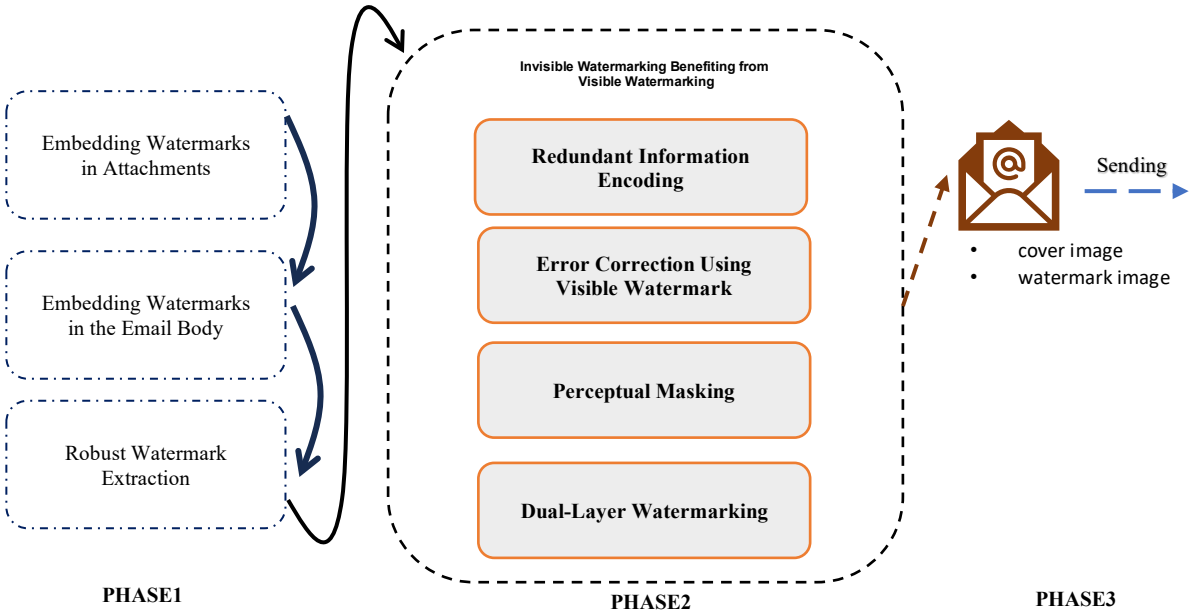



**Fig. 1** *Proposed framework for sender steps*

In PHASE 1, the embedding watermark in the cover Image converts the watermark image to grayscale and normalizes its pixel values to the range [0, 1]. Subsequently, generate a transparent layer by multiplying the normalized watermark by the transparency factor Alpha to create a faint, almost invisible watermark. To embed a watermark, a transparent watermark layer needs to be added to cover the image as illustrated in Algorithm 1, blending it into each color channel (R, G, B) and clipping pixel values to ensure they remain within the valid range [0, 255]. To generate the Watermarked Image that appears visually similar to the original image and contains an embedded watermark as output of this step. In Phase 2, Redundant Information Encoding ensures that repeating parts of the watermark are encoded, allowing for recovery of any lost or corrupted parts from the other parts. The Error Correction Using Visible Watermark: A visible watermark acts as a guide or helper. If the invisible part is corrupted, the visible part can fix it. In Perceptual Masking, the process of hiding the watermark in less noticeable parts of the image is mixed with other parts. The final step in PHASE 2, Dual-Layer Watermarking, uses both visible and invisible watermarks together. The visible one can serve as a warning or proof of ownership, while the invisible one provides security or verification background. Where algorithm 1 describes the procedure of the sending side.

**Algorithm** 1: sender email with Watermark

**Input:** Cover Image; Watermark Image

**Output:** email with Watermark attachment

1:Read the Cover_Image
2:Read Watermark_Image.
3:Resize the Watermark_Image to match the dimensions of the Cover
4: **for** each Watermark[k] in the Cover Image **do**
5: *Watermark[x] =Convert Watermark[x] to grayscale*
*6: Normalized Watermark[k]= Normalize Watermark[x] pixel values to the range [0, 1].*
*7: Transparent layer[k] = Normalized Watermark[k]*
*8: Embed Watermark =Transparent layer[k]+ Cover Image*
*9:* ***End for***

10: //Embedding Invisible Markers in the Email Content:
11:Create Invisible Pattern[k] embedding as watermarks in attachments.
12:Invisible Watermark Integration.
13:Start Sending.
**End.**

During the step of embedding invisible markers in email content, it is necessary to generate an invisible pattern by creating subtle formatting changes (such as spacing, capitalization, and special characters) in the Email_Content. This involves embedding invisible markers into the email's text to create an invisible watermark that is undetectable to the human eye but recognizable to a detection algorithm. A protected email containing the embedded invisible markers should have a watermarked image attached and include the protected email as the message body, then be sent to the intended recipient in PHASE3. At the receiving end, the watermark should be reconstructed based on a known Alpha value using a reversible technique from the Watermarked_Image received. Additionally, ensure that the extracted watermark matches the original one; if the watermark is missing or altered, flag the email as potentially compromised. Detection of invisible markers is performed by analyzing the protected email content for embedded markers. If the markers are missing or modified, trigger a warning of possible tampering.

Finally, perform a Performance Metrics Evaluation by calculating PSNR (Peak Signal-to-Noise Ratio) to assess the visual quality of the Watermarked_Image and compute BER (Bit Error Rate) to determine the integrity of the extracted watermark. Evaluate NC (Normalized Correlation) to measure the similarity between the original watermark and the derived watermark. Use these metrics to classify emails as either secure or compromised. This approach generates alerts and logs if any watermarking or invisible markers have been tampered with; the logs assist with detection results for audit purposes, providing evidence for any potential cybersecurity incident.

This methodology aims to improve email security and increase the detectability of any unauthorized tampering or data theft attempts. Additionally, the proposed approach aims to enhance the robustness of invisible watermarking by leveraging the benefits of visible watermarking. The proposed approach's mechanism is implemented through several steps, as outlined in Algorithm 1. Embedding watermarks in attachments by employing a hybrid watermarking technique, which involves placing an apparent and inconspicuous watermark on all the attached documents or images. Both types of watermark, such as a name or an invented field (email of the recipient, transaction ID, or time-stamped hash), can make the signature almost invisible. This makes it a visible watermark that identifies the sender and recipient of the email, thus enhancing the tracking of unauthorized data transfers in attachments. Hybrid watermarking is also an effective countermeasure to phishing as it assigns a specific code to official templates from organizations. Whenever these emails come from trusted senders, automated systems can check the watermark and promptly identify the impersonators, labeling such spoofs as illustrated in Fig. 2.

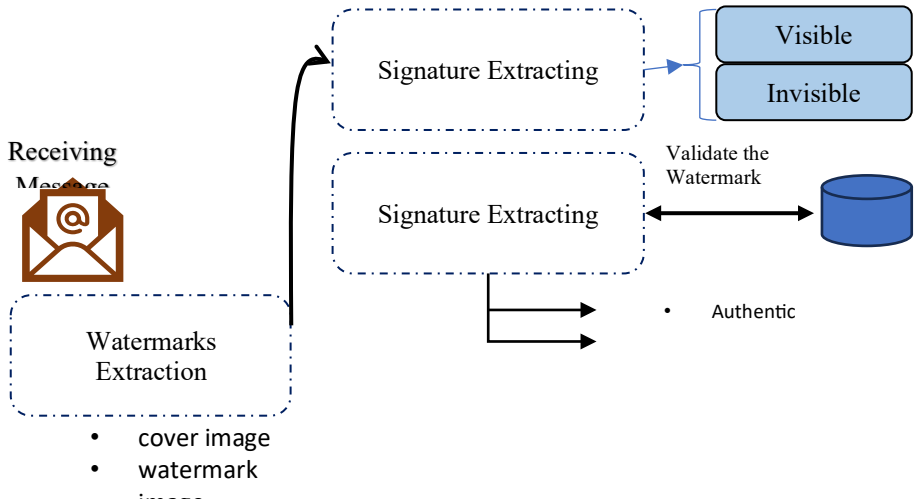



**Fig. 2** *Examine receive to the detected unsolicited incident*

Systems can verify the embedded watermark, swiftly identifying any attempts to impersonate and flagging these spoofed emails as shown in Fig 2.

Robust watermark extraction by the receiver, a cyberattack performed, or any data leak detected. The watermark could be retrieved from the email body or an attachment, where an unauthorized user is most likely to lurk. This extraction process assists in the identification of the first receiver or last legitimate user in tracing out the unauthorized access point described in Algorithm 2.

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 435-447

**Algorithm** 2: Receives email with Watermark

**Input:** email with Watermark attachment
**Output:** Cover Image; Watermark Image

1. read email content
2. read the attachment content
3. ***for each*** content **do**
4. *extract watermark[k] by :*
5. *Watermark[k] = Redundant Information Encoding*
6. *Watermark[k] = Error Correction Using Visible Watermark*
7. *Watermark[k] = Perceptual Masking*
8. *Watermark[k] = Dual-Layer Watermarking*
9. ***end for***
10. **End.**

This type of extraction exploits invisible and visible watermarks by performing several ways as follows:

- Redundant Information Encoding: one of the uses of the visible watermark is to provide additional information with respect to the invisible watermark, for example, if part of the invisible watermark were erased for some reason, such as an attack or compression, then the system can restore the lacking information on the basis of the visible watermark.
- Error Correction Using Visible Watermark: A visible watermark can include an error correction code that supports an invisible watermark. While noise or compression degrades the hidden watermark, the clearly visible watermark enables one to recapture or verify the information.
- Perceptual Masking: Significant information that can be of importance when designing the embedding of invisible watermarks can be drawn from visible watermarking, for instance, visible watermark distortions change areas that are not very conspicuous to the naked eye. These same regions can also be used to place invisible watermarks, thereby increasing robustness, as the area covered by the watermark is altered in a way that makes it difficult to notice or attack.
- Dual-Layer Watermarking: The first watermark that is detected can be an ink drop watermark, which can be embedded on top of a second invisible watermark. This could help detect visible watermark tampering and automatically influence the invisible one, providing a means of easily identifying tampering or such modifications.

In this approach, we performed an analysis on two prevalent invisible and visible watermarking methods. The first one, LSB (Least Significant Bit), operates in the spatial domain, while the second one, DWT (Discrete Wavelet Transform), operates in the frequency domain to provide a balanced analysis. In this paper, the network's performance metrics, such as throughput, jitter, packet delivery ratio, and latency, are not examined. For visible watermarking, our approach involves three components: A carrier cover file, a serial number, and a specific Algorithm. This method can be mathematically expressed with the following equations 1 and 2 [17]:

$$Img_{cw} = \alpha 1 Img_c + \alpha 2 Img_w \quad (1)$$

$$Img_w = \frac{Img_{cw} - \alpha 1 Img_c}{\alpha 2} \quad (2)$$

Where, $Img_c$ is the cover image, $Img_w$ is a logo watermark, and $Img_{cw}$ is watermarked image. The parameters α1 and α2 play a critical role in determining the visibility and concealment of the watermark image. These values are between zero and one and should be chosen to optimize the values so that their sum equals one at any given time, ensuring the watermark is not so prominent as to lose its significance.

The three files in this procedure are the cover image file, the logo image file, and the watermarked image file. The parameters serve as the guardians of the watermark's invisibility, carefully balancing their values between zero and one to ensure that their combined strength always adds up to one.

The invisible watermarking technique, derived from the Least Significant Bit (LSB) method, inevitably embeds the watermark signal into the least significant areas of the cover image. As the quality of the cover file fluctuates but on a minimal level, such changes are insignificant and so unnoticeable that they fit right in with

what we may call unnecessary information. This method is capable of achieving high load and is nearly invisible with respect to a target; however, it lacks resilience. However, it offers a higher level of performance compared to the LSB technique.

## 4. System Implementation and Evaluation

RGB Lena image BMP format bean size is 512 x 512 pixels is selected as cover image, where a string of characters serial number was adopted as invisible watermark image, bean size is also 512 x 512 bits. The simulation was done on a personal computer that was used with an advanced Core i7-2.10 GHz, 16GB RAM, with a 4GB RTX GPU, machine coded in Python. Moreover, using Simple Mail Transfer Protocol or SMTP. It is a communication standard for sending electronic mail from one node to another using a network, notably the Internet. In contrast, sending, receiving, or forwarding email, mail servers, or the transfer agents known as message transfer agents use SMTP.

Three watermarking techniques were used to embed the watermark into the cover image: An LSB, a DCT, and the proposed method that this study aims to examine. However, it was here where things only started to get better. It was then subjected to a round of vigorous tests for its endurance, scaling, restructuring, adding noise, and removing noise. First, in the scaling attack, the watermarked image was reduced in size from 512 x 512 pixels to just 256 x 256 before being enlarged again in order to check its flexibility. The next test was the reformat attack, where the efficiency of the file in transition from BMP to JPG and vice versa was tested. The noising attack introduced a subtle but dangerous 0.1% Gaussian noise similar to the sound of static in the background. Last of all came the de-noising attack, and the watermarking had to survive the exquisite clean up using the Wiener filter (to remove noise caused by corrupted signal), which entailed both spatial and frequency domains. Every time it tested my watermark to the maximum.

The integrity and quality of the watermarked image were assessed using PSNR measurements, both before and after these attacks, to evaluate just how strong and enduring the watermark truly was under these extreme conditions. equation (3) and equation (4)[14].

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

$$MSE = \frac{\sum_0^i \sum_0^j \sqrt{Img_o - Img_w}}{ij} \tag{4}$$

Where MSE stands for the mean square error. Where Imgo and Imgw are the original and watermarked image, respectively, of dimension (ixj) pixels.

In other words, Normalized Correlation (NC) is essential for assessing the degree of an invisible watermark in binary form. These metrics explain the extent to which distortion or confusion happens when the watermark is under attack. The level of disruption is quantified in the following equations. The extent of disruption is captured mathematically through the following equation[17]:

$$BER = \sum_0^j \sum_0^j (w_{0+} w_R))/ij \tag{5}$$

$$NC = \sum_0^j \sum_0^j (w_0 \ . \ w_R))/ \ \sum_0^j \sum_0^j (w_0)^2)^{0.5} \ (\sum_0^i \sum_0^j (w_R)^2)^{0.5} \ \ldots\ldots 6 \tag{6}$$

where $w_0$ represent the original watermark file and $w_R$ represent the reconstructed watermark file, both are in binary format with a size of (ixj) pixels. When applying the proposed technique, a watermark was embedded into the carrier file quite subtly, having an invisibility ratio of only 0.3%.

The watermarked file was then copied cautiously through email. Subsequently, the level of invisibility was increased to 5%, challenging the real effectiveness of the technique and its robustness. This made it possible to compare its stability when using it with LSB and DCT, known sophisticated methods that this subtle but rather complex approach can offer.

The summary of the Algorithm is essentially to overlay a subtle, semi-transparent watermark over the attachments of the email, overlaying invisible tags within the body of the email. Set a template that contains the watermarked content and send the email with it. Extract and analyze the watermark and invisible markers at the recipient's side. Identify unsupervised changes by using PSNR, BER, and NC methods. If tampering is predicted, the user is notified and writes a log for the tampering incident.

It has been observed that, on average, the proposed method achieves acceptable image quality and robustness of the watermark. The Mean Square Error (MSE) of 21 is relatively high and corresponds to a Peak Signal-to-Noise Ratio (PSNR) of 20.66 dB, indicating a noticeable decrease in the watermarked image quality compared to the original. The high BER of 0.9678 suggests that many watermark bits have been altered or missed, which could impact the watermark extraction process to a certain extent. Nevertheless, despite some differences,

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 435-447

the Normalized Correlation (NC) value of 0.8812 shows that the extracted watermark and the original are still quite similar, clearly indicating that the watermark remains reasonably recognizable even at this stage, with a balance between robustness and fidelity as shown in Fig 3.


**Fig. 3** *Original cover image and watermarked image*

In this research, both the quantitative and qualitative data were selected to enable precise and easily analyzable respondent outcomes. The impact of the watermarked file's quality on PSNR (dB) before and after exposure to scaling, reformatting, noise addition, and denoising attacks, with an invisibility ratio of 3%, is quantitatively shown in Table 2. It can be clearly observed that, for high-quality watermarked files, the proposed technique is slightly inferior to LSB and DWT. However, after performing these attacks, the proposed method mitigates the impact of the attacks. Tables 3 and 4 illustrate the comparison between the extracted watermark's BER and NC as measured by LSB, DWT, and the proposed method before and after performing the attacks on the sample image. It is desirable for BER to approach zero and for NC to approach one. Tables 3 and 4 clearly illustrate that the proposed technique yields the lowest BER and highest NC compared to other techniques, with DWT slightly outperforming LSB. Moreover, Table 5 contains the author's qualitative analysis of Table 2, while Tables 6 provide a qualitative analysis of Tables 3 and 4. Lastly, Table 7 shows that the watermarked file, along with its quality and reliability, remains intact before and after sending it through the network as an email, with an invisibility ratio of 0.3%. Table 7 demonstrates the superiority of the proposed technique over other methods, which are considered less reliable than techniques such as DWT and LSB.

Table 5 demonstrates the potential of the proposed system by showing that the cover image remains unaffected by attacks. In Table 6, we observe that the extracted watermarks are not significantly impacted by attacks targeting files, since the proposed model operates at the network application layer and behind a firewall, which filters received files and extracts watermark files. This allows the system to function efficiently in terms of email security.

**Table 2** *PSNR shows that the watermarked files' quality stays stable before and after several attacks using various insertion methods*

| | | With Attack | | | |
|---|---|---|---|---|---|
| **Tech.** | **No Attack** | **Scale** | **Reformat** | **Nois** | **Denoise** |
| LSB | 30.99 | 28.49 | 26.71 | 27.45 | 23.17 |
| DWT | 50.18 | 32.22 | 28.69 | 29.99 | 26.54 |
| Proposed | 50.98 | 32.24 | 28.69 | 30.05 | 23.96 |

**Table 3** *The BER of a watermark file extracted using various insertion techniques is measured before and after different attacks as a function of time*

| Tech. | No Attack | With Attack | | | |
|---|---|---|---|---|---|
| | | Scale | Reformat | Nois | Denoise |
| LSB | None | 0.18 | 0.39 | 0.40 | 0.50 |
| DWT | None | 0.38 | 0.44 | 0.39 | 0.48 |
| Proposed | None | 0.11 | 0.22 | 0.01 | 0.9 |

**Table 4** *Examine the extracted watermark file's NC before and after injection attacks*

| Tech. | No Attack | With Attack | | | |
|---|---|---|---|---|---|
| | | Scale | Reformat | Nois | Denoise |
| LSB | None | 0.20 | 0.48 | 0.60 | 0.38 |
| DWT | None | 0.47 | 0.39 | 0.69 | 0.61 |
| Proposed | None | 0.79 | 0.66 | 0.99 | 0.94 |

**Table 5** *The quality of watermarked files using various insertion techniques before and after exposure to different attacks*

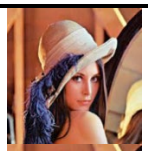
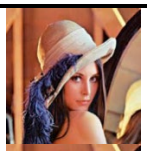
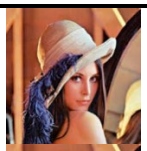
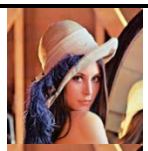
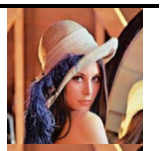
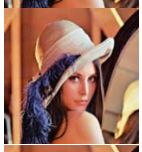
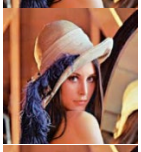
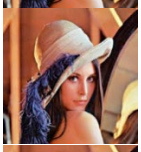
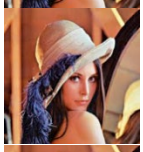
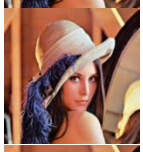
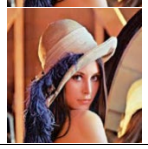
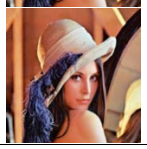
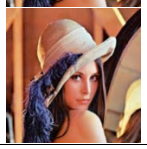
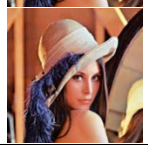
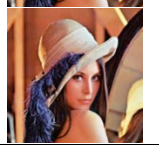
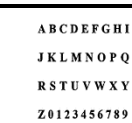
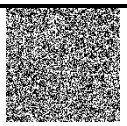
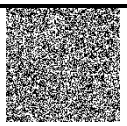
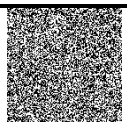
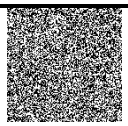
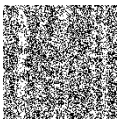
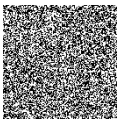
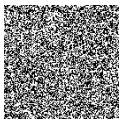
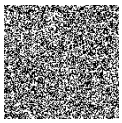
| Tech. | Before Attacks | After Attacks | | | |
|---|---|---|---|---|---|
| | | Scale | Reformat | Nois | Denoise |
| LSB |  |  |  |  |  |
| DWT |  |  |  |  |  |
| Proposed Method |  |  |  |  |  |

**Table 6** *The extracted watermark file using various insertion techniques was assessed before and after exposure to different attacks in terms of subjective measurement*

| Tech. | Before Attacks | After Attacks | | | |
|---|---|---|---|---|---|
| | | Scale | Reformat | Nois | Denoise |
| LSB |  |  |  |  |  |
| DWT |  |  |  |  |  |
| Proposed Method |  |  |  |  |  |

J. of Soft Computing and Data Mining Vol. 6 No. 1 (2025) p. 435-447

**Table 7** *Watermarked file quality and extracted watermark file resilience in subjective and objective measures before and after email transmission through the network*

| Tech. used through Sending over Email | PSNR | SerialNo. | BER | NC |
|---|---|---|---|---|
| LSB | 20.66 | ABCDEFGHI<br>JKLMNOPQ<br>RSTUVWXY<br>Z0123456789 | 0.9678 | 0.8812 |
| DWT | 20.66 | ABCDEFGHI<br>JKLMNOPQ<br>RSTUVWXY<br>Z0123456789 | 0.9678 | 0.8812 |
| Proposed Method | 76.7 | ABCDEFGHI<br>JKLMNOPQ<br>RSTUVWXY<br>Z0123456789 | 0.98.23 | 0.8937 |

The results demonstrate the significant performance of the proposed approach under various types of cyber attacks, including scaling, reformatting, noise addition, and denoising attacks, as it employs hybrid watermarking (HWM), which enhances security by combining multiple techniques, such as spatial and frequency domain methods. The proposed method increases robustness against common attacks. As a result, the watermark is harder to detect or remove, offering stronger protection for digital content in emails.

## 5. Conclusion

The methodology presented herein provides a flexible and effective way to enhance email security by employing watermarking techniques. This approach makes it virtually impossible for attackers to modify and manipulate email content without being detected. By combining both visible and invisible watermarking, the proposed methodology achieves a synergistic effect: visible watermark, in addition to augmenting the effectiveness of the invisible one, constitutes an unmistakable signal against any tampering with the watermark. This makes the system highly robust against various types of attacks, such as compression, tampering, or cropping. The visible watermark serves as both a protector and a protection measure, enhancing the invisible watermark and making it particularly useful in applications that require enhanced security or the prevention of tampering with related content. Furthermore, the flexibility of the proposed framework can be further enhanced by fine-tuning all the embedding, extraction, and error-correction algorithms to achieve the desired level of invisibility. The findings show that the proposed method is completely secure against a variety of threats, including scaling, reformation, addition of noise, and removal of noise, far exceeding the performance of traditional methods that work in either the spatial domain (LSB method) or the frequency domain (such as the DWT method). Moreover, the watermarks extracted using the proposed method are also easily readable and distortion-free, as compared to the watermarks obtained by other approaches, which can be completely distorted. Notably, the quality of the watermarked files is generally good and competes equally well with offline methods. As for the future, the combination of encryption with the proposed approach could certainly enhance its security characteristics even further. Through the integration of visible and invisible watermarking approaches in the proposed methodology, security is complemented by the quick identification of unauthorized changes, which are detected through alerts generated in response. Another enhanced feature is tracing capabilities, which enable watermarks to relate emails and attachments to individual users for improved tracking and reporting. The proposed approach characterizes a sound and preventive mechanism for the current problems of email protection.

## Acknowledgement

The authors would like to thank the University of Mosul, College of Computer Sciences and Mathematics for their provided facilities, which helped to improve the quality of this work.

## Conflict of Interest

The authors declare no financial, personal, or professional conflicts of interest that could bias this research's conduct, results, or interpretation. The authors have no conflicts of interest related to this work.

## Author Contribution

*The authors confirm contribution to the paper as follows:* ***study conception and design:*** *Ibrahim M. Ahmed, Awos Kh. Ali;* ***data collection:*** *Ibrahim M. Ahmed;* ***analysis and interpretation of results:*** *Ibrahim M. Ahmed, Mahmood S. Mahmood;* ***draft manuscript preparation:*** *Ibrahim M. Ahmed, Awos Kh. Ali. All authors reviewed the results and approved the final version of the manuscript.*

## References


[1] Al-Dabbagh, M., Ali, A. K., & Alhasan, A. L. S. (2022). Employing light fidelity technology in health monitoring system. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), 26(2), 989-997, *http://doi.org/10.11591/ijeecs.v26.i2.pp989-997.*

[2] Alnaish, Z. A. H., & Hasoon, S. O. (2023). Hybrid binary whale optimization algorithm based on taper shaped transfer function for software defect prediction. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska, 13, *https://doi.org/10.35784/iapgos.4569.*

[3] Younus, Z. S., & Alanezi, M. (2023). Detect and Mitigate Cyberattacks Using SIEM. 2023 16th International Conference on Developments in eSystems Engineering (DeSE), *https://doi.org/10.1109/DeSE60595.2023.10469387.*

[4] Tayachi, M., Nana, L., Pascu, A. C., & Benzarti, F. (2023). A hybrid watermarking approach for DICOM images security. Applied Sciences, 13(10), 6132, *https://doi.org/10.3390/app13106132.*

[5] Ahmed, I. M., & Kashmoola, M. Y. (2022). Investigated insider and outsider attacks on the federated learning systems. 2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), https://doi.org/10.1109/COMNETSAT56033.2022.9994330.

[6] Hajjaji, M. A., Gafsi, M., Ben Abdelali, A., & Mtibaa, A. (2019). FPGA implementation of digital images watermarking system based on discrete Haar wavelet transform. Security and Communication Networks, 2019(1), 1294267, https://doi.org/10.1155/2019/1294267.

[7] Devi, K. J., Singh, P., Dash, J. K., Thakkar, H. K., Santamaría, J., Krishna, M. V. J., & Romero-Manchado, A. (2022). A new robust and secure 3-level digital image watermarking method based on G-BAT hybrid Optimization. Mathematics, 10(16), 3015,https://doi.org/10.3390/math10163015.

[8] Nam, S.-H., Yu, I.-J., Mun, S.-M., Kim, D., & Ahn, W. (2020). WAN: Watermarking attack network. arXiv preprint arXiv:2008.06255, https://doi.org/10.48550/arXiv.2008.06255.

[9] Lidyawati, L., Darlis, A. R., Jambola, L., & Kristiana, L. (2022). Digital watermarking image using three-level discrete wavelet transform under attacking noise. Bulletin of Electrical Engineering and Informatics, 11(1), 231-238, https://doi.org/10.11591/eei.v11i1.3565.

[10] Fatahbeygi, A., & Tab, F. A. (2019). A highly robust and secure image watermarking based on classification and visual cryptography. Journal of information security and applications, 45, 71-78, https://doi.org/10.1016/j.jisa.2019.01.005.

[11] Dai, Z., Lian, C., He, Z., Jiang, H., & Wang, Y. (2022). A novel hybrid reversible-zero watermarking scheme to protect medical image. IEEE Access, 10, 58005-58016, https://doi.org/10.1109/ACCESS.2022.3170030.

[12] Mohammed, A. O., Hussein, H. I., Mstafa, R. J., & Abdulazeez, A. M. (2023). A blind and robust color image watermarking scheme based on DCT and DWT domains. Multimedia Tools and Applications, 82(21), 32855-32881, https://doi.org/10.1007/s11042-023-14797-0.

[13] Yang, P., Lao, Y., & Li, P. (2021). Robust watermarking for deep neural networks via bi-level optimization. Proceedings of the IEEE/CVF international conference on computer vision, https://doi.org/10.1109/ICCV48922.2021.01457.

[14] Aslantaş, P. V., Khorsheed, F. H., & Ahmed, B. (2020). Wrapper feature selection approach based on binary firefly algorithm for spam e-mail filtering. Journal of Soft Computing and Data Mining, 1(2), 44-52. https://doi.org/10.30880/jscdm.2020.01.02.005.

[15] Gao, G., Wang, M., & Wu, B. (2023). Efficient Robust Reversible Watermarking Based on ZMs and Integer Wavelet Transform. IEEE Transactions on Industrial Informatics, https://doi.org/10.1109/TII.2023.3321101.

[16] Radha Kumari, R., Vijaya Kumar, V., & Rama Naidu, K. (2024). Deep learning-based image watermarking technique with hybrid DWT-SVD. The Imaging Science Journal, 72(6), 749-765, https://doi.org/10.1080/13682199.2023.2223080.

[17] Abd, M. H., & Allawi, O. W. (2023). Secured Mechanism Towards Integrity of Digital Images Using DWT, DCT, LSB and Watermarking Integrations. Ibn AL-Haitham Journal For Pure and Applied Sciences, 36(2), 454-468, https://doi.org/doi.org/10.30526/36.2.3088.

[18] Kadian, P., Arora, S. M., & Arora, N. (2021). Robust digital watermarking techniques for copyright protection of digital data: A survey. Wireless Personal Communications, 118, 3225-3249. https://doi.org/10.1007/s11277-021-08177-w.

[19] Rofiatunnajah, N. B., & Barmawi, A. M. (2023). Improving ANiTW Performance Using Bigrams Character Encoding and Identity-Based Signature. IEEE Access, 11, 24257-24280. https://doi.org/10.1109/ACCESS.2023.3254586.