

Fragile Watermarks Detecting Forged Images

Hala K. Hussein^{*1}, Ra'ad A. Muhajjar², Bashar S. Mahdi³

¹ Computer Science Dept., College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq

² Computer Science Dept., College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq

³ Computer Science Dept., College of Computer Science, University of Technology, Baghdad, Iraq,

Correspondence

Hala k. Hussein

Department of Computer science,

College of Computer Science and Information Technology,

University of Basrah, Basrah, Iraq.

Email: itpg.hala.khalid@uobasrah.edu.iq

Abstract

Technology and digital communications have advanced so that digital photos, videos, or text may be easily manipulated by those not authorized to do so. In addition, the availability of specialized picture editing programs like Photoshop has simplified the process of altering photographs. At first glance, there may seem to be no problem, especially when an image editing method is necessary to delete or add a certain scene that increases the picture's beauty. But what about personal images or images with copyright? Attempts are constantly made to spoof these images using different approaches. Therefore, measures to reduce the likelihood of counterfeiting in digital and printed forms of media are required. The proposed approach aims to detect a counterfeit in images using a unique generator that conceals the data represented by the embedded watermark utilizing modern visual cryptography and hash algorithms. Image extractions may easily be analyzed for signs of forgery. As a result, our approach will detect and validate phony documents and images.

KEYWORDS: Fragile watermark, spatial domain, LSB, Image.

I. INTRODUCTION

The richness and complexity of multimedia information pose significant challenges to human life and activity. One of the most widely shared forms of multimedia on the web is digital images, which can be easily copied and edited. Users thus have a significant issue in securing the data sent via it [1][2]. In recent years, protecting sensitive data has grown more crucial. Advancements necessitate new methods of data transmission security in transmission technologies. Watermarking is a kind of security system that has been developed as a technological solution for information security [3]. The word "watermark" refers to a group of bits used to identify the private information added to an image to prevent unauthorized usage. For maximum protection, the watermark should be integrated into the image rather than distinct from it. The image may be entirely undetectable to human sight while still being readable by computers, and the quality of the image is kept with little loss [4]. Two domains in which the watermark may be implemented are the spatial domain and the Transform domain. Among the several methods of digital watermarking, spatial domain watermarking is the most straightforward. Spatial domain watermarking has a lengthy history. While developing the embedding and extraction algorithms, researchers offered approaches to adding the watermark to the original picture by modifying the pixel values in the spatial domain, similar

to methods reported in previous publications using this strategy. While spatial watermarks are easy and fast to install, their fragility causes them to be seen as weak. Multiplicative watermarking techniques, including the more well-known Transform-domain watermarks, are widely considered secure against attacks [5]. Digital watermarking might be visible or invisible, but our work would be invisible. Additionally, a watermark may be robust, fragile, or semi-fragile [6].

II. RELATED WORK

Raj & Shreelekshmi, 2018, they utilized two fragile watermarks in this paper as additional security. They started by using MD5 to build a 128-bit representation of the original picture by dividing it into 8x8 chunks. The first fragile watermark is represented as a 2LSB and is embedded in each block. The second method is identical to the first, except that the image is broken up into 16x16 blocks. Use of the SHA-256 hashing method is at the core of the watermark's generation. The produced 256 bits are saved in the cover image's least significant bit [7].

Chitra, K., & Prasanna Venkatesan, V., 2018, VC with watermarking was suggested as a technique. The picture's colors have been converted to binary, and the resulting image has been divided in two. As a result of this method, the receiver will get a cover image that includes one of the



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. Published by Iraqi Journal for Electrical and Electronic Engineering by College of Engineering, University of Basrah.

shares. The watermark must be embedded before it can be extracted for matching. Thus, extraction is performed by the receiver, and OR, or XOR is employed during stacking. According to the results, XOR rather than OR increases PSNR in this study[8].

Ayu et al., 2019, the suggested method offered dual-layer fragile digital watermarking, which involves two watermarks being incorporated into medical images. They used a method known as the Advanced Encryption Standard. (AES) to protect the secrecy and validity of the Electronic Patient Record (EPR). Digital Imaging and Communication in Medicine (DICOM) tags were embedded in 2LSB as the first fragile watermark, and their integrity was then verified using the Secure Hash Algorithm (SHA256). Finally, they divided the picture into non-overlapping chunks, gave each one an id, computed the SHA256 for integrity, and used it as a tamper detector [9].

Gul & Ozturk., 2019, LSB was proposed as a method for spatial watermarking. They made a watermark by breaking an image into four sub-blocks and then processing the three sub-blocks through the SHA-256 hash method. Checking the watermark's authenticity and integrity is possible by extracting and comparing it to a hash value. Due to this, the authentication procedure will be less secure, particularly against pixel-based attacks such as salt and paper noise [10]. Sinhal et al., 2020 the color picture was separated into 2*4 non-overlapping blocks using pseudo-random, created a random integer between 0 and 1 and then converted to binary. The output of this procedure is a 6-bit sequence. It would then be a fragile watermark embedded. Block-wise classification as forgery or originally based on neighbor blocks is performed in the extraction watermark by performing a series of operations such as LSB and others to each RGB (red, green, blue) channel. Utilize the six most critical bits in recovery [11].

Reyes-Reyes et al., 2021, the author suggests a system for dealing with RGB color pictures by dividing them into non-overlapping blocks. After that, use Pseudocode on each channel to create the recovery and authentication watermarks for each block. In the second phase, one bit of authentication data for blocks is generated by performing the XOR operation on each of the three recovery watermarks [12].

Su et al., 2021. Proposed hidden information using the concept of the Sudoku game. After producing it using a pseudo-random number generator, they embedded fragile watermarking. During extraction, compare it to the original. They employed a two-layer hidden information basis from the first layer to extract nine candidate pairings. After that, extract the block index and values index from each pair. They chose the final pair with the watermark based on Euclidian distance [13].

III. ARCHITECTURE OF PROPOSED SYSTEM

The suggested type is built on a three-stage design, the first stage defining the generated watermark. Step two involves adding a watermark, which is done in a certain technique. The last step involves the receiver extracting the watermark from the image and then verifying the integrity of the embedded watermark. This method relies on a multi-layer

approach to accomplish its goal at each stage. Figure 1 shows the general framework of the suggested approach:

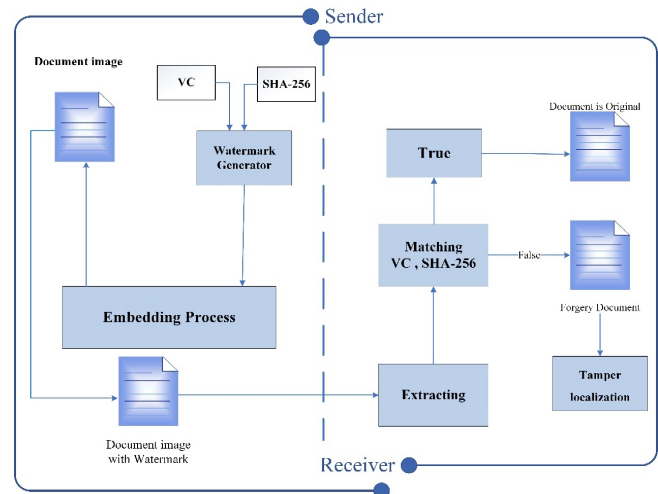


Fig. 1: Architecture of proposed system.

A. Stage one (Watermark Generator)

A unique generator using visual cryptography and hashing is employed to secure our images from this vulnerability. By ignoring redundant bits, watermark generators minimize the number of bits concealed in the cover image document.

1) Visual Cryptography (VC)

The proposed technique employs Visual Cryptography (VC) to generate two random images from a secret image, one for private sharing and one for public sharing, in order to guarantee the best level of security and authenticity for the owner's watermark. A private sharing image is used in the creation of the watermark. In addition, a secret one is kept to be utilized in the watermark extraction process and to ensure that an attacker cannot uncover any hints about a concealed picture in the original images. The proposed visual cryptography method entails creating the following step:

- Distribution phase secret encode table
- Reconstruction phase [14]

This approach assumes that an image or message is composed of black pixels represented as (0) and white pixels represented as (1). A combination of share stacking and an OR operation revealed the secret. In other words, the probability of obtaining a white (or black) pixel is 50 percent [15].

2) Hash Function

Hashing is Converting random data into a fixed-length representation (a hash code or message digest) that is often represented as a hexadecimal number. The hash function uses an index to get the information that corresponds to the original value or key. One-way hash functions are the only function used to determine the hash value from a given input without the ability to recover the initial information. The SHA-2 hash method is implemented. While it is an excellent method for keeping your data safe, it does take a while to process. A 256-bit digest is generated using SHA-2[16].

B. Embedding stage (Sender)

This stage is often performed by the sender, responsible for protecting the picture against tampering by entering sensitive data by a specified plan with the recipient. The Two tasks are within the purview of the sender: implementing VC and using a hash (SHA256). These two processes, as described before, generate the secret bits.

1) Pre-Processing System

The suggested method is designed to operate with color images. A high-quality digital image was obtained in this instance. The first step in this procedure is ensuring that the image size is 900 pixels wide by 1250 pixels tall. As a result, the sender obtains the image and proceeds to the next system step.

2) Divided Blocks

Using the watermarking generator, the sender splits the image into eight blocks of the same size, embedding a certain amount of confidential data. The pixels inside the selected block will be explicitly set. The red, green, and blue (RGB) values for each pixel in a color image are individually kept in an array of sizes M by N by 3, where M is rows and N is columns. First, the RGB color image has been separated into individual layers. Then chosen a pixel is, its color's integer value is transformed to binary. Further, the proposed embedding approach benefits from parallel processing of the three-band color image.

3) Pixel Per Pixel

A specific pixel has been chosen at this point. The primary action is to split the pixel color image into three layers. After selecting a pixel, the integer value for color is converted to binary. At this point, the locations of the pixels that will be embedded will be determined. Selecting pixels one by one is the standard method for finding altered areas.

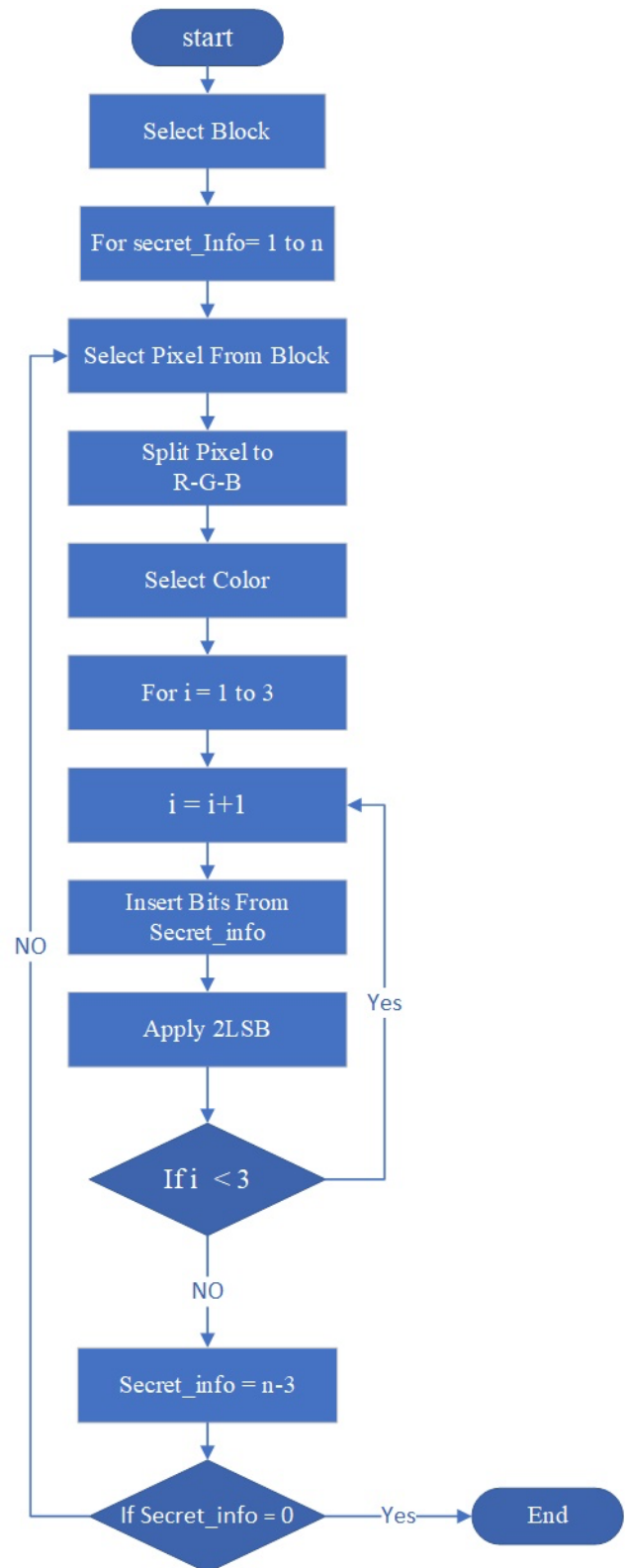
4) Embedding Process

The embedding procedure starts with selecting a block and the required pixels inside that block. Each color in this picture includes 8 bits after converting the pixel to binary RGB images. Also, the pixel has two parts—the most significant bit (MSB) and the least significant bit (LSB), each of which is four bits. LSB enables the system to operate in the spatial realm by changing pixel values without decreasing picture quality. The key benefit of LSB is that it has essential features: (Ease of implementation, writing code, execution time, and storage requirements) since it trades the value of another location without needing complex calculations. The following flowchart1 explains the embedding mechanism within a single block, and these steps are repeated on the eight blocks until all confidential information is embedded.

C. Extracting stage (Receiver)

The extraction process is based on the block and pixel requirements established during the embedding phase. Separating the image into eight blocks and giving each pixel one of the three RGB colors allows easy extraction of the watermark bits. In most cases, there are two steps to the extraction procedure. The first step must recover eight hash codes before the VC bits can be extracted and the logo recreated. The extraction process is the opposite of the

embedding process. The extraction process is the opposite of the embedding process.



Flowchart 1: The process of embedding in one block.

IV. IMPLEMENTED MEASURES

A. Mean square error (MSE)

It shows the exact square difference between the un-watermarked and watermarked versions of the image. In general, MSE has no precise value, although the lower the value, the better, and zero are optimal. The MSE define as equation1[17]:

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M \times N} (C_i - S_i)^2 \quad (1)$$

while $M \times N$ is the size of the image, C_i and S_i are, respectively, watermark and host images. a main problem with MSE it depends to intensity of pixel.

B. Imperceptibility

Regarding watermarks, imperceptibility is one of the most important aspects to consider. As a result, the amount of noise introduced to the picture must be quantified by the watermark bits to evaluate the quality of the reconstructed image. This unit of measurement is referred to as an image quality metric. The Peak Signal to Noise Ratio (PSNR) define as equation (2) [18]

$$PSNR = 20 \times \log_{10}(MAXI) - 10 \times \log_{10}(MSE) \quad (2)$$

Max is the highest pixel value that may be achieved, while MSE is the mean square error. .it important to know that increasing the ratio of PSNR will improve image quality, which is the most important factor in the proposed system


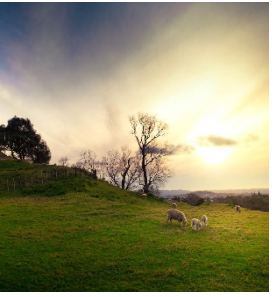

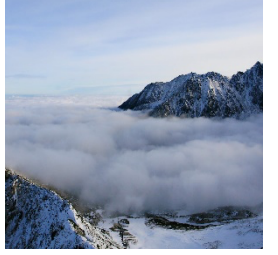
V. RESULTS OF EXPERIMENT

The supplied form has been executed on a dataset taken from the Kaggle website as part of this proposal [19]. Table I represents this dataset and the result after embedding the watermark with the measures. Figure 2 shows the original image and the watermarked image, and Figs. 3,4,5,6 will include models with a histogram for the original and embedded image drawn by python 3.9. As shown, there is no difference between the histogram of the original image and after the embedding process, which proves that the image quality was not affected by inserting the watermark. And that is the power of our work. The fragile watermark was employed using an earlier technique for generating the secret bits to authenticate the image and identify any manipulation.



(a) Original image (b) watermarked image
Fig. 2: TEST1 original image and the watermarked image

TABLE I
THE RESULT AFTER AN ATTACK ON AN IMAGE WITH A FRAGILE WATERMARK.

 <p>TEST1 image PSNR=70.83db MSE =0.003</p>	 <p>TEST2 image PSNR=69.03db MSE =0.008</p>
 <p>TEST3 image PSNR=69.08db MSE =0.008</p>	 <p>TEST4 image PSNR=67.26db MSE =0.01</p>

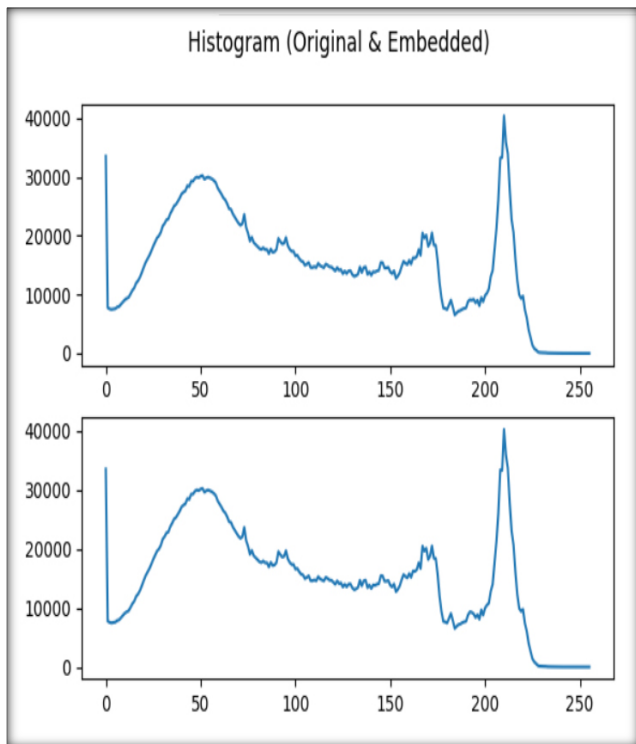


Fig. .3: show a histogram for the original TEST1 image and the watermarked image after embedding.

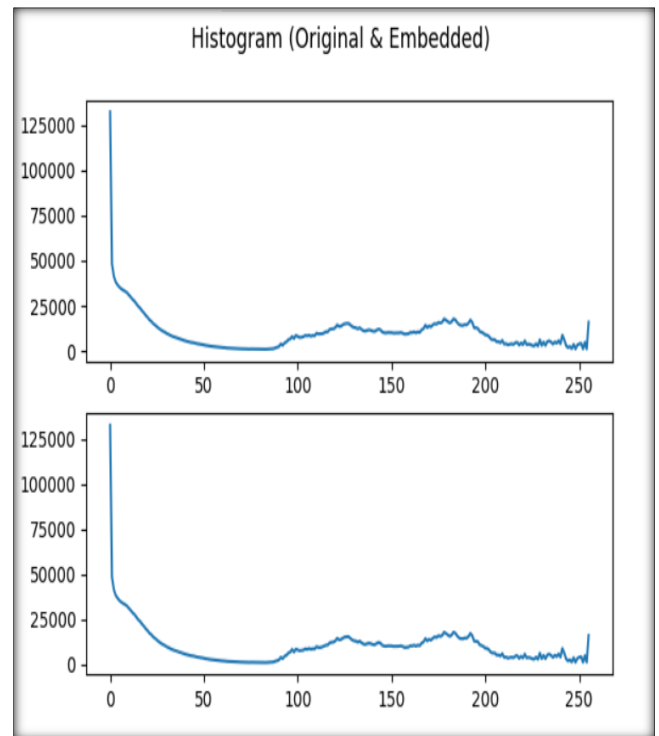


Fig. .5: show a histogram for the original TEST2 image and the watermarked image after embedding.

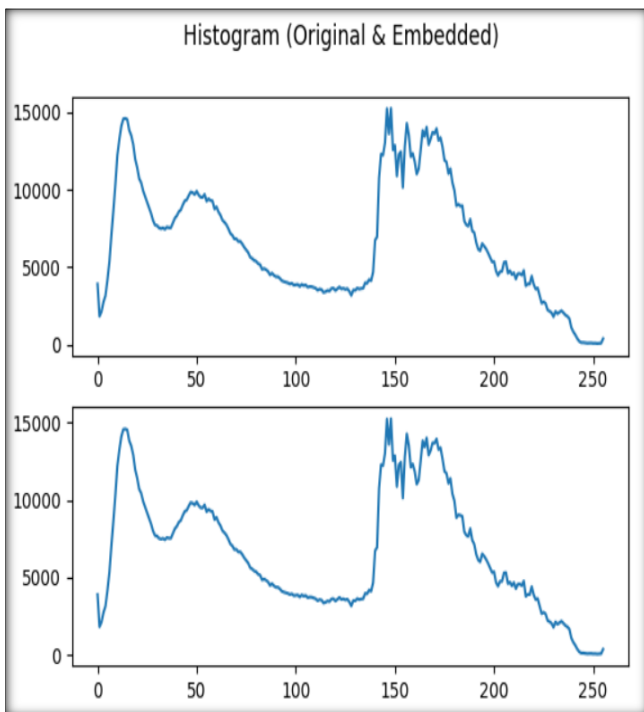


Fig. .4: show a histogram for the original TEST3 image and the watermarked image after embedding.

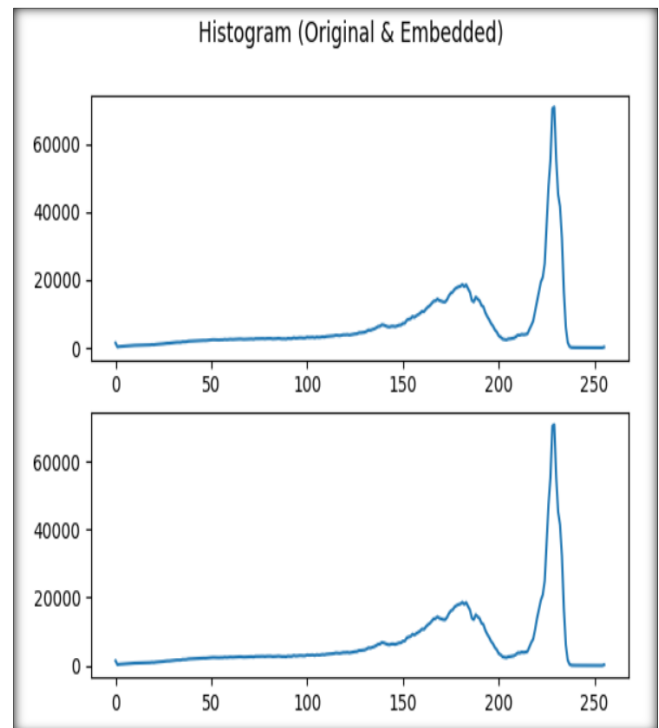


Fig. .6: show a histogram for the original TEST4 image and the watermarked image after embedding.

The major purpose of the suggested framework is not only to construct a watermark and embed it inside an image, but this mark must also be fragile to identify any manipulation that happens on the image, which means that it will be destroyed once an alteration occurs. This manipulation is considered forging if it isn't within the acceptable range.

The VC serves as a threshold for determining whether or not an image is fraudulent. This VC is created by combining two photos to ensure that the authentication is genuine and not fabricated. Additionally, the hash function's benefits have been used by giving one-way outputs, providing that this procedure had greater integrity.

After conducting the embedding procedure on several images, this technique yielded the best PSNR scale results. However, the quality of the photograph remained intact. This study found that the PSNR value increased with picture size, with a maximum value of about 70 dB and a minimum of 59.15, and a reduction in MSE of 0.07 as image size was increased, as shown in Table II

TABLE II
IMAGE RESULTS AFTER DIFFERENT ATTACKS


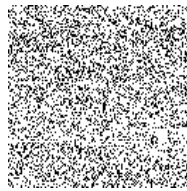
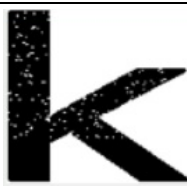
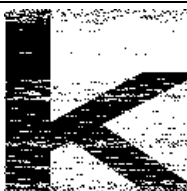




Size	PSNR	MSE
512.512	59.15	0.07
800.800	63.05	0.03
1000.1500	66.75	0.013
1600.2400	70.82	0.005

The VC serves as a threshold for determining whether or not an image is fraudulent. This VC is created by combining two photos to ensure that the authentication is genuine and not fabricated. Additionally, the hash function's benefits have been used by giving one-way outputs, providing that this procedure had greater integrity. Some image assaults have been performed to validate the suggested system's operation and influence the image's quality. Implemented the attacks (Rotate, salt and pepper, scale).

Measurements are shown in the Table III below. Following the attacks, the image was rotated and checked whether the original bits could be recovered once they returned to their original shape. It was able to overcome this assault. On the contrary, in scale attacks, there was no way to extract the watermark because it was too fragile, and the pixel values changed when zoomed in or out, affecting the extraction process. In regards to the Salt and Pepper assault, the watermark was successful, and the majority of the bits were recovered.

Another attack involved the deletion of an image, and a component from another image was copied and pasted into the image where the watermark was embedded. As indicated in the table, the suggested work successfully detected this manipulation process. The effects were visible in the logo through apparent distortion, which helped determine the location of the alteration.

TABLE III
IMAGE RESULTS AFTER DIFFERENT ATTACKS

Lena image	Attacks	PSNR	MSE	LOGO EXTRACTED
	Rotate (90, -90)	70.83 DB	0.005	
	Scale	48.13 DB	1	
	Salt and Pepper	49.81 DB	0.67	
	Write on image (more the 500 word)	34.65 DB	22.24	
	compression	70.82 DB	0.005	
	Gaussian filter	39.64 DB	7.05	
	Delete from image	-	---	
	Copy a component from one image to another	---	---	

VI. A COMPARISON OF THE SEVERAL ALTERNATIVE SYSTEMS

The proposed method was compared to previous techniques to determine if image quality is maintained after embedding. Suggested results showed a higher percentage of PSNR compared to other methods, showing the technique's ability to preserve image quality. as shown in Table IV and Fig. 7 explain the proposed method with other techniques.

TABLE IV
SHOWS A COMPARISON WITH OTHER METHODS

Method	PSNR (db)
Raj & Shreelekshmi, [7]	55.14
Chitra, K., & Prasanna Venkatesan, V., [8]	56.49
Ayu et al. [9]	44.91
Gul & Ozturk [10]	57.19
Sinhal et al., [11]	49.68
Reyes-Reyes et al. [12]	43.89
Su et al [13]	47.8
Proposed Method	70.82

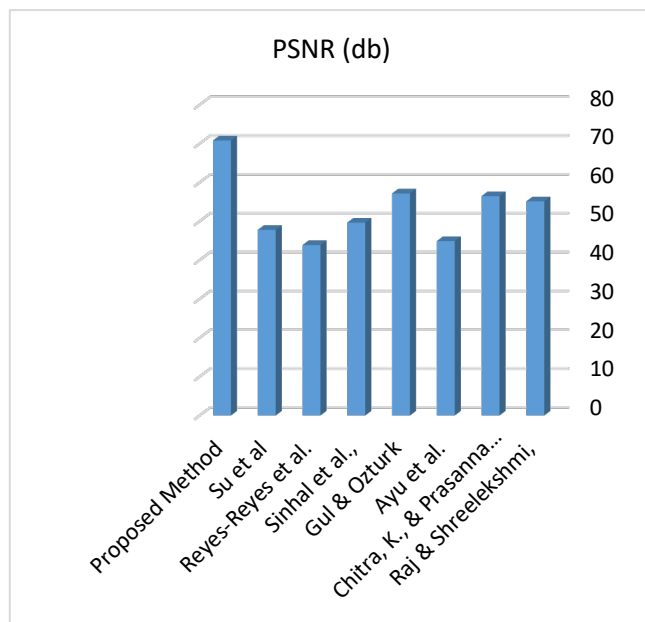


Fig. 7: Comparison of PSNR chart with other methods.

VII. CONCLUSION

This research study proposed a method for creating a watermark and embedding it in images to verify their authenticity or fraud. The watermark has shown its fragility in several manipulations. It has successfully overcome various attacks on an image, such as noise that typically happens to the picture for different reasons. Additionally, the embedding procedure in the second bit contributed to the

extra image fragility required to accomplish the study's purpose. Furthermore, the use of lossless compression in the form of PNGs assisted in maintaining the quality of the embedded images, which achieved a PSNR of 70.82db and an MSE of 0.005; both values are considered to be within acceptable ranges for picture quality.

This research helped create a new fragile watermark generator using hashing and visual cryptography techniques, providing an additional layer of protection against watermark tampering by ensuring that the attacker cannot determine the nature of the watermark being entered. As a result, the technique employed in this research may be used for more sensitive images and documents, such as government files, to detect any alteration on these documents and the watermark's inherent fragility in this methodology. The suggested approach is very efficient at preserving image quality.

CONFLICT OF INTEREST

The authors have no conflict of relevant interest to this article.

REFERENCES

- [1] K. Dou, B. Guo, and L. Kuang, "A privacy-preserving multimedia recommendation in the context of social network based on weighted noise injection," *Multimedia Tools and Applications*, vol. 78, no. 19. pp. 26907–26926, 2019. doi: 10.1007/s11042-017-4352-3.
- [2] A. Alsimry, K. Hussein Ali, and E. Wahab Abood, "A new approach for finding duplicated words in scanned Arabic documents based on OCR and SURF," *J. Basrah Res.*, no. 47, 2021. [Online]. Available: <https://www.iasj.net/iasj/journal/260/issues>.
- [3] Farah Abdul-Hussain Badr, "secure data communications using cryptography and IPv6 steganography.pdf," *International Journal of Engineering & Technology*, vol. 7, no. 4.19, pp. 624-628, 2018.
- [4] H. M. Abdul-Nabi, E. S. Al-Shawi, and H. L. Hussain, "Hiding Three Images at one image by Using Wavelet Coefficients at Color Image," *Basrah Journal of Science (A)*, vol. 28, no. 1, pp. 58-73, 2010.
- [5] Q. Su et al., "New Rapid and Robust Color Image Watermarking Technique in Spatial Domain," *IEEE Access*, vol. 7, pp. 30398–30409, 2019. doi: 10.1109/ACCESS.2019.2895062.
- [6] D. Singh and S. K. Singh, "DCT based efficient fragile watermarking scheme for image authentication and restoration," *Multimedia Tools and Applications*, vol. 76, no. 1. pp. 953–977, 2017. doi: 10.1007/s11042-015-3010-x.
- [7] N. R. N. Raj and R. Shreelekshmi, "Blockwise Fragile Watermarking Schemes for Tamper Localization in Digital Images," *2018 International CET Conference on Control, Communication, and Computing, IC4 2018*, pp. 441–446, 2018. doi: 10.1109/CETIC4.2018.8530950.

- [8] K. Chitra and V. Prasanna, "A Dynamic Security Model for Visual Cryptography and Digital Watermarking," *Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018)*, pp 351–357, 2018.
- [9] M. A. Ayu, T. Mantoro, and I. M. A. Priyatna, "Advanced watermarking technique to improve medical images' security," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 17, no. 5. pp. 2684–2696, 2019. doi: 10.12928/TELKOMNIKA.v17i5.13292.
- [10] E. Gul and S. Ozturk, "A novel hash function based fragile watermarking method for image integrity," *Multimedia Tools and Applications*, vol. 78, no. 13. pp. 17701–17718, 2019. doi: 10.1007/s11042-018-7084-0.
- [11] R. Sinhal, I. A. Ansari, and C. W. Ahn, "Blind Image Watermarking for Localization and Restoration of Color Images," *IEEE Access*, vol. 8, pp. 200157–200169, 2020. doi: 10.1109/ACCESS.2020.3035428.
- [12] R. Reyes-Reyes, C. Cruz-Ramos, V. Ponomaryov, B. P. Garcia-Salgado, and J. Molina-Garcia, "Color image self-recovery and tampering detection scheme based on fragile watermarking with high recovery capability," *Applied Sciences (Switzerland)*, vol. 11, no. 7. 2021. doi: 10.3390/app11073187.
- [13] G. D. Su, C. C. Chang, and C. C. Chen, "A hybrid-Sudoku based fragile watermarking scheme for image tampering detection," *Multimedia Tools and Applications*, vol. 80, no. 8. pp. 12881–12903, 2021. doi: 10.1007/s11042-020-10451-1.
- [14] S. A. Alsuhbany, "Developing a Visual Cryptography Tool for Arabic Text," *IEEE Access*, vol. 7, pp. 76573–76579, 2019. doi: 10.1109/ACCESS.2019.2920858.
- [15] S. A. Fadhil, A. K. Farhan, and A. H. Radie, "Visual Cryptography Techniques: Short Survey," *4th Int. Iraqi Conf. Eng. Technol. Their Appl. IICETA 2021*, pp. 276–282, 2021, doi: 10.1109/IICETA51758.2021.9717352.
- [16] N. Hamed and A. Yassin, "Secure Patient Authentication Scheme in the Healthcare System Using Symmetric Encryption," *Iraqi Journal for Electrical and Electronic Engineering*, vol. 18, no. 1. pp. 71–81, 2022. doi: 10.37917/ijeeec.18.1.9.
- [17] O. Evsutin, A. Melman, and R. Meshcheryakov, "Digital steganography and watermarking for digital images: A review of current research directions," *IEEE Access*, vol. 8. pp. 166589–166611, 2020. doi: 10.1109/ACCESS.2020.3022779.
- [18] L. Rakhmawati, Wirawan, and Suwadi, "Image Fragile Watermarking with Two Authentication Components for Tamper Detection and Recovery," *2018 Int. Conf. Intell. Auton. Syst. ICoIAS 2018*, pp. 35–38, 2018. doi: 10.1109/ICoIAS.2018.8494080.
- [19] <https://www.kaggle.com/datasets/pavansanagapati/images-dataset>