ISSN 2007-9737

Secure Medical Image Authentication Using Zero-Watermarking based on Deep Learning Context Encoder

Rodrigo Eduardo Arevalo-Ancona, Manuel Cedillo-Hernandez*,
Ana Elena Ramirez-Rodriguez, Mariko Nakano-Miyatake,
Hector Perez-Meana

Instituto Politécnico Nacional, SEPI,
Mexico

[rarevaloa0900, aramirezr0906]@alumno.ipn.mx,
[mcedilloh, mnakano, hmperezm]@ipn.mx

Abstract. Zero-watermarking is a robust and lossless technique for digital image security, copyright protection, and content authentication. This paper introduces a novel zero-watermarking scheme for medical image authentication based on deep learning. The proposed approach leverages a neural network based on the Context Encoder to extract distinctive features from the image, enhancing the method. The training of the neural model increases robustness. The watermark consists of a halftone image of the patient's face, serving as a unique identifier for medical study. By revealing the watermark, medical professionals can verify the correspondence between the imaging study and the patient. Therefore, an XOR operation merges the watermark sequence and the extracted features. The proposed method offers continuous image protection, safeguarding sensitive medical data. Extensive experiments demonstrate the technique's robustness against various attacks, including geometric transformations (scaling, cropping, resizing, rotation) and image processing manipulations (filtering, blurring, JPEG compression, and noise addition). The detection watermark process achieves a low bit error rate and a high normalized cross-correlation, validating the method's robustness and effectiveness. The deep neural network improved the robustness of the presented zero-watermarking scheme making it suitable for practical applications in medical data security and integrity.

Keywords. Zero-watermarking, image security, image, authentication, deep learning, feature extraction.

1 Introduction

In recent years, the exponential growth in digital content use and distribution has increased with the development of digital technology. Digital images have facilitated services in the health area, such as telemedicine, e-learning, or remote assistance [1]. In addition, healthcare systems provide easy access to medical data, which could be manipulated or redistributed without authorization. Consequently, Medical imaging requires security, patient data privacy, and diagnostic accuracy [2], [3]. In this context, one of the principal challenges is the need for technology developments for image authentication, protection, and security.

Watermarking techniques are used for copyright protection, image authentication, and image protection [4]. Traditional watermarking techniques embed a signal into the image generating a distortion [5, 6]. Image distortion is not suitable for scenarios where the image integrity must be preserved. Arum Patel and Prabhat Patel proposed a novel hybrid watermark algorithm in their work, [7] which leverages wavelet coefficients for color components adapting.

They utilized the Singular Value Decomposition (SVD) technique on the LL and HH sub-bands obtained from the 2nd level of the Discrete Wavelet Transform (DWT) from the logo image. In another study by Sinhal and Ansari [8], a dual watermark scheme was designed to preserve the regions of interest in medical images.

First, a robust watermark is embedded into the image using the Integer Wavelet Transform (IWT). Simultaneously, to ensure the integrity of the regions of non-interest (RONI), a fragile watermark was incorporated by replacing the least significant bit (LSB).

doi: 10.13053/CyS-28-1-4898

During the detection stage, a deep neural network is deployed to extract the watermarks effectively. In their research [9], Qasim, Meziane, and Aspin proposed an innovative reversible and imperceptible watermarking scheme for detecting distortions on magnetic resonance images.

To minimize image distortion run-length encoding is used to compress the watermark. During the embedding process, the image undergoes segmentation for ROI detection, achieved through histogram thresholding. By identifying smooth blocks inside the ROI of the medical image that matches those in the watermark image.

Zero-watermarking has emerged as an innovative and effective technology for image authentication. Zero-watermarking schemes are a lossless technique for digital image security, copyright protection, and content authentication. Zero-watermarking hides information into a created stego-image or master share, merging features from the host image and the user’s watermark sequence [10]. Zero-watermarking systems do not embed information into the image, keeping its quality intact without distortions.

One of the advantages of zero-watermarking lies in its non-intrusive nature. Unlike other methods that directly embed information into the image, zero-watermarking preserves the image's quality without introducing any distortions. This aspect is particularly critical in medical imaging, where maintaining the integrity of the original image is paramount for accurate diagnosis and analysis. Tayachi et al. [11] developed a hybrid watermark algorithm designed to protect DICOM images. The image is partitioned into two regions: ROI (Region of Interest) and RONI (Region of Non-Interest) using a thresholding technique.

In the zero-watermarking scheme, features from the ROI area are combined with the watermark. The non-zero-watermarking algorithm embeds multiple copies of the watermark in the RONI area. The embedding process employs a linear interpolation technique. Hosny and Darwish introduced an innovative zero-watermarking technique specifically designed for color medical images, as described in their research.

To enhance the security of the watermark, in [10] they applied the Arnold transform, effectively scrambling the watermark to prevent unauthorized access or tampering. The image features are obtained using the multi-channel fractional-order Gegenbauer moments (FrMGMs) of color images, which represents unique features of each image.

The master share construction combines the scrambled watermark with the extracted image features using the XOR logic operation. Rocek et al. introduced a reversible watermarking technique for medical imaging in their research [12]. They divided the image into ROI and RONI areas since the ROI areas cannot be distorted.

The zero-watermarking process employed the Dual-Tree Complex Wavelet Transform (DT-CWT) on the ROI. From the DT-CWT, they selected the LL coefficients to extract essential image features, which were combined with the watermark. To enhance the security, a reversible contrast mapping technique for RONI watermarking is applied.

Additionally, deep learning has revolutionized various fields, including computer vision, image processing, and image security. Deep learning techniques in zero-watermarking have shown promising results. Deep neural models help in the detection process of more robust and discriminative features.

Deep neural networks for watermarking systems enhance the security and resilience of the authentication process. Fierro-Radilla et al. generate a zero-watermarking using convolutional neural networks [13]. The authors presented an architecture centered around convolutional neural networks (CNN) comprising 13 layers and one fully connected layer for feature extraction.

This process allows the creation of a matrix containing the most relevant features extracted from the image. Subsequently, an XOR logic operation is applied to effectively merge the feature matrix with the watermark. Gong et al. introduced in their research a robust zero-watermarking for medical images based on the DenseNet model [14].

The DenseNet neural network is employed, and a dense block is added to obtain a residual block that facilitates the extraction of crucial feature maps related to the image. These feature maps were associated with the watermark. Furthermore, to enhance the security of the watermark, the authors implemented an encryption process using a logistic map algorithm.

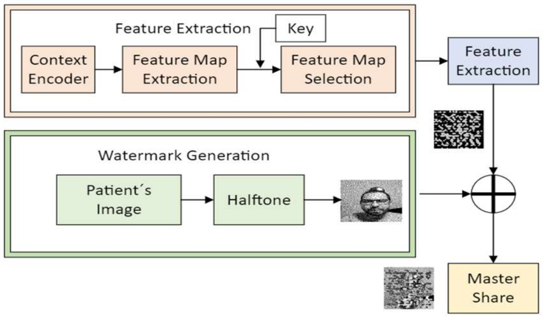


Fig. 1. Zero-watermarking encryption diagram

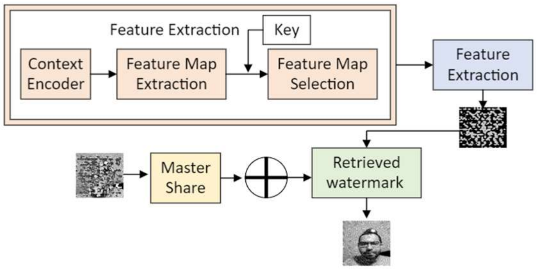


Fig. 2. Zero-watermarking diagram for watermark retrieval

Huang et al. presented a robust zero-watermarking system based on VGG network for healthcare information security [15]. A chaotic scrambling is applied for the encryption of the watermark, ensuring heightened security.

Next, the feature map is obtained using the VGG pre-trained model. Thus, it is combined with the image features to construct the hash. As a result, the hash becomes intricately linked with the scrambled watermark. Han et. al used the VGG19 neural network model to propose a zero-watermarking scheme for medical images [16].

The authors used the VGG19 model to extract deep feature maps directly related to the original medical image. These feature maps were fused to generate a comprehensive feature image. Therefore, it is applied the Discrete Fourier Transform to this image to obtain a feature matrix, which is combined with the watermark. To enhance the security of the watermark.

Several challenges associated with zero-watermarking methods can be summarized as follows: 1. Limited focus on specific image types: Many existing methods tend to concentrate on a particular type of image, which might limit their applicability to diverse image datasets and scenarios. 2. A narrow focus on specific attacks: Certain methods were designed to be robust against to specific attacks, which may leave the watermark vulnerable to other potential tampering. 3. Usage of small watermarks: Some proposed techniques employ small-sized watermarks, which might compromise the watermark's robustness and visibility in certain situations. 4. Overreliance on logo watermarks: The majority of utilized watermarks are often simple logo images, which might lack the capacity to embed complex information or provide sufficient security against sophisticated attacks.

The main contribution of this paper lies in the feature extraction using a Context Encoder, which significantly enhances the system's robustness. The Context Encoder is a deep neural model that learns specific features to facilitate watermark recovery, even in cases where the watermark has been tampered. The Context Encoder was designed for image reconstruction according to its context background.

Although the approach primarily focuses on medical images, the experiments also conduct tests with natural images, demonstrating excellent performance in image protection. This versatility allows the system to safeguard different types of images, beyond its initial scope. Furthermore, the watermark size is 160x160 pixels, constituting a halftone image derived from the patient's face.

This approach enables the treating doctor to determine if the study belongs to the patient. In the case of natural images or photographs, allow the author recognition to protect the image copyright. This additional level of detail and personalization enhances the overall security and authentication capabilities of the watermarking system.

In summary, the Context Encoder's application for feature extraction for watermarking methods, the adaptability to different image types, and the utilization of specific watermarks contribute to the paper's significant contributions in advancing image protection and authentication techniques.

The rest of the paper is organized as follows: In Section 2, the proposed method is comprehensively explained, detailing the innovative approach used for zero-watermarking. Section 3 presents the experimental results obtained from the conducted tests, showcasing the performance and effectiveness of the proposed

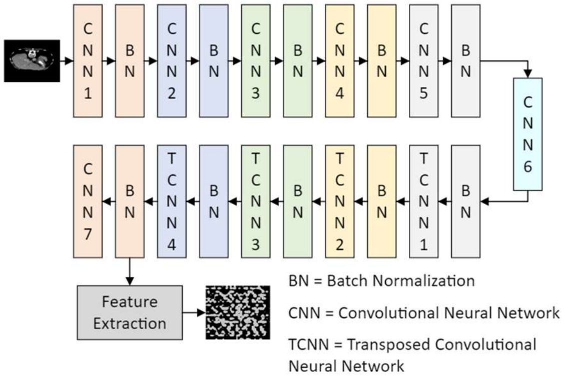


Fig. 3. Neural network architecture

Table 1. Neural network hyperparameters

Layer	Neurons	Stride	Padding	Kernel Size
CNN 1	64	2	1	4
CNN 2	128	2	1	4
CNN 3	512	2	1	4
CNN 4	256	2	1	4
CNN 5	512	2	1	4
CNN 6	1024	2	1	4
TCNN 1	512	2	1	4
TCNN 2	256	2	1	4
TCNN 3	128	2	1	4
TCNN 4 (Feature Map Extraction)	64	2	1	4
CNN 7	3	2	1	4

method. Finally, in Section 4, the conclusions drawn from this research are presented, summarizing the key findings, and highlighting the significance of the proposed approach in the context of image protection and watermarking.

2 Zero-Watermarking Proposed Method

This paper introduces a robust zero-watermarking scheme specifically designed for medical images. However, it shows robustness in natural images. The feature extraction stage is based on the Context Encoder, a generative deep learning model used for inpainting. Moreover, the watermark is a halftone image derived from the patient's face.

On the other hand, the watermark employed in this scheme is larger in size (160x160 pixels) compared to most of the watermarks used in previous research. The Context Encoder learns unique features from the image. This neural network model increases the robustness of the zero-watermarking system against different tampering attacks. The general diagram of the zero-watermarking technique is depicted in Fig. 1 and Fig. 2, providing a visual representation of the algorithm.

The algorithm is detailed in the following subsections. Context Encode model: This section explains the architecture of the Context Encoder for feature extraction. Watermark Generation: Describes the process of generating the watermark, ensuring it is suitable for embedding within medical and natural images. Master share generation: Illustrates the encryption method used to enhance the security of the watermark, ensuring its protection against geometric and advanced image processing attacks. Watermark Retrieval: Explains the procedure for retrieving the watermark, allowing verification and authentication purposes. The overview of the proposed zero-watermarking technique, this paper contributes to image protection and authentication in medical and natural image domains.

2.1 Context Encoder Model

Deep learning is a branch of machine learning that utilizes data to learn features through pattern recognition, enabling more informed decision-making. Deep learning algorithms aim to perform tasks more efficiently, resulting in better learning of extracted characteristics from the data [17].

One area where deep learning has made significant strides is image feature extraction. It enables automatic learning and discrimination of features from the data. In deep neural networks, lower layers obtain simple features like edges and textures, while deeper layers extract more complex and abstract features, such as high-level patterns.

This hierarchical feature extraction process allows the neural network to learn representations and structures from the image data, leading to a better understanding and recognition of visual patterns. The implementation of deep learning models for feature extraction has notably improved the performance of various computer vision tasks [18, 19] surpassing traditional approaches.

The Context Encoder [20], is specifically designed for inpainting techniques. The model focuses on approximating the likelihood of pixels to create new data samples using a generator. A discriminator is employed to test the similarity of the generated data to the original dataset through a probability distribution.

The generator (*G(z)*) utilizes features from the dataset to create new data, while the discriminator (*D(z)*) distinguishes the differences between the generated samples and the original dataset (1) [18]:

$$\min_{G}\max_{D} V(D,G) = E_{x\sim P(data)}[LogD(x)] + E_{x\sim P(z)}[Log(1-D(G(z))], \tag{1}$$

where the loss function of the discriminator log(D(z)) learns the features from the dataset using the probability from the similarity between the dataset image and the reconstructed $(x{\sim}P(data))$ image to compare the loss function of the generator log(1-D(G(z))) and approximates the equilibrium between the features of the dataset and the new samples $(x{\sim}P(z))$.

This paper focus on the Context Encoder, due to its ability to learn unique features. These features are specific to each image, contributing to the robustness of the watermarking process.

The training process is conducted using different geometric and image advanced processing attacks on the medical images, ensuring the model can identify only the most crucial features relevant to watermarking.

The incorporation of such a refined feature selection approach facilitates the creation of a secure master share, ensuring the protection of sensitive patient information embedded in the watermark sequence. Fig. 3 provides a clear illustration of the neural network model's architecture.

Overall, our emphasis on the Context Encoder and the strategic training of the neural network result in an effective and robust watermarking technique, ensuring authentication and integrity in medical image applications.

The neural network employs the rectified linear activation function (ReLU) on each layer specifically for region of interest detection, effectively identifying essential features crucial for constructing the master share. The feature map is extracted from the TCNN 4 layer (Transposed Convolutional Neural Network), subsequently, the selected feature map is binarized.

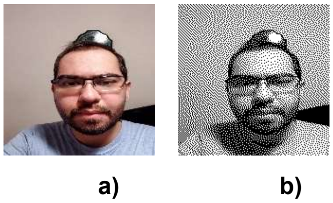


Fig. 4. a) Original patient's image, b) Halftone version of patient´s image

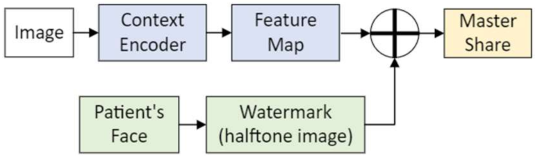


Fig. 5. Master share construction

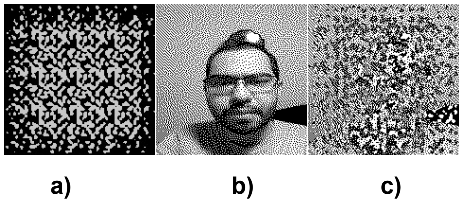


Fig. 6. a) Image features, b) Halftone image watermark, c) Master share

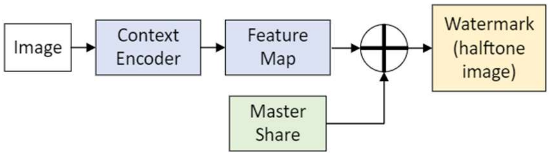


Fig. 7. Watermark retrieval

The Mean Squared Error (MSE) defined by (2) is used as the loss function. The MSE computes the average error and quantifies the distance between the image feature values (x) and the extracted features $(\overline{x})$. A lower error value indicates that the predictions are closer to the actual feature values [21]:

$$MSE = \frac{1}{n}\sum(x-\bar{x})^2. \tag{2}$$

Table 1 summarizes the configuration of the neural network's layers, specifying the number of neurons, kernel size, and stride.

A ReLU activation function is applied for each layer and MSE loss, the neural network effectively identifies and captures the significant image features, allowing for accurate and reliable construction of the master share. Finally, the feature maps are obtained from the TCNN. The selection of the specific feature map is accomplished using a key, which allows us to determine and extract the desired features from the network.

This selective feature map extraction process ensures that only relevant and critical features are used for the master share construction, enhancing the efficiency of the zero-watermarking algorithm. With the key-based selection, we can focus on the most important features needed for our zero-watermarking system, enabling better performance.

2.2 Watermark Generation

The watermark utilized in this technique is created from an image of the patient's face, which is transformed into a halftone image. Halftoning is a widely used image processing technique that converts continuous-tone grayscale images into binary halftone representations.

The main objective of halftoning is to replicate the appearance of various shades of gray by strategically distributing black and white pixels, thereby simulating the illusion of grayscale tones through carefully placed dots of varying sizes. This process ensures that the watermark effectively captures the visual information of the original image while being suitable for embedding and authentication purposes [22]. The halftone image (*Ih*) transform is calculated for each pixel from the gray image (*Ig*) using (3):

$$Ih(i,j) = Ig(i,j) + \sum h(m,n)e(i-m, j-n)\,, \qquad (3)$$

where $h(m,n)$ in (4) is the error filter and $e(i - m, j - n)$ in (5) is the quantization error:

$$h(m,n) = \begin{bmatrix} 0 & Ig(i,j) & 7 \\ 3 & 5 & 1 \end{bmatrix}, \qquad (4)$$

$$e = u(i,j) - Q, \qquad (5)$$

where *Q* is the binarize pixel value (6):

$$Q = \begin{cases} 0 & if \quad Ig(i,j) < threshold \\ \quad 1 & if \quad other \end{cases}. \qquad (6)$$

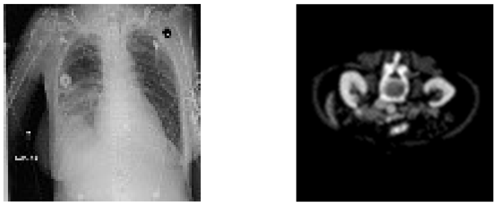
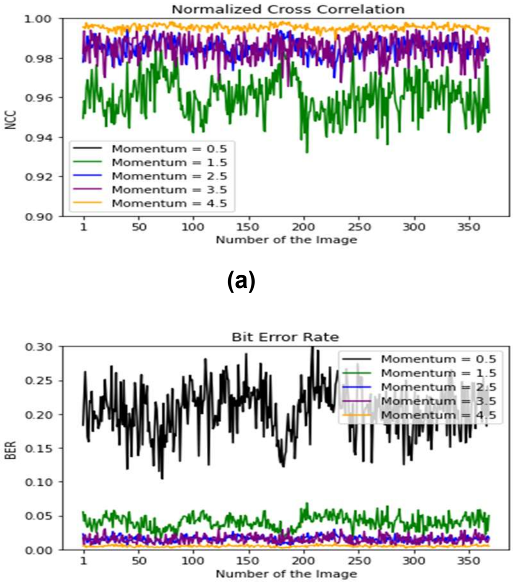


Fig. 8. Test medical images



(a)

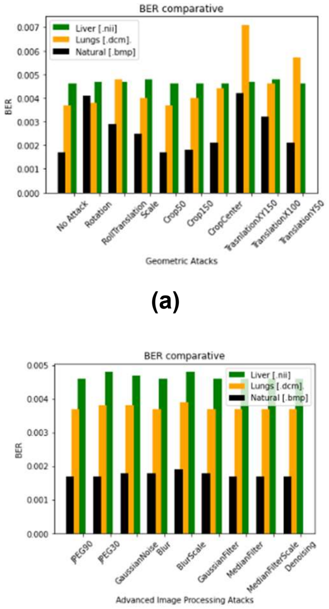
(b)

Fig. 9. a) Momentum BER, b) Momentum NCC

Fig. 4a displays the original image of the patient, while Fig. 4b illustrates the same image after undergoing the conversion process to a halftone representation. The halftone technique ensures the preservation of key visual features of the patient's image. In addition, it is suitable as watermark for the master share construction authentication purposes.

2.3 Master Share Generation

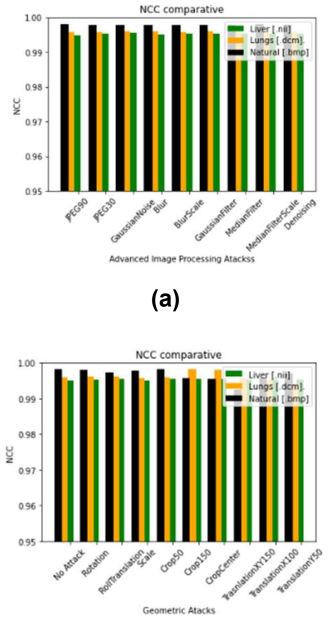
The master share plays a crucial role in image authentication and certification, which is store securely in an external device [23]. The master share (*MS*) consists of a stego-image that incorporates information from the host image.



(a)

(b)

Fig. 10. a) Geometric attacks, b) Common signal processing



(a)

(b)

Fig. 11. a) Geometric attacks, b) Common signal processing

This information is derived from the combination of extracted features (*ef*) through the Context Encoder, and the binarized watermark (*W*) by an XOR (⊕) logic operation given by (7):

$$MS = ef \oplus W. \tag{7}$$

This process can be visualized in the diagram presented in Fig. 5. In Fig. 6, we present a graphical representation of the extracted features from the image, as well as the watermark and master share generation process.

2.4 Watermark Retrieval

The watermark retrieval (*W'*) validates the patient's identity and certifies the ownership of the imaging study. The watermark is revealed by (8), which involves combining the master share (*MS*) with the extracted features (*ef*) obtained from the pre-trained Context Encoder.

This process allows us to securely reveal, verify the watermark, and confirm the ownership of the medical data:

$$W' = ef \oplus MS. \tag{8}$$

This process is in the diagram from the Fig. 7.

This process verifies the integrity and authenticity of the imaging study, providing a robust watermarking system for medical image applications.

3 Experiments Results

The experiments were conducted to evaluate the robustness and efficiency of the proposed algorithm. The main testing image database [24] is comprised with 123 liver images, each with a size of 512 x 512 in Neuroimaging Informatics Technology Initiative (NII) format.

However, to increase the diversity of the dataset, the image base was expanded to 369 images, capturing various perspectives of the same liver images.

Moreover, the method's robustness was tested using 369 images with 512 x 512 size in Digital Imaging and Communications in Medicine (DICOM) format in modality x-ray obtained from [25]. An Example of test medical images are shown in Fig. 8. Additionally, the algorithm was validated with 369 natural images (.bmp) from [26]. This extensive testing ensured the algorithm's versatility and suitability for various image types. The watermark consists of a halftone image derived from the patient's face, sized at 160 x 160 pixels.

This personalized approach reinforces the security and uniqueness of the watermark, making it highly effective for image authentication and ownership certification purposes. Additionally, data augmentation techniques were applied to enhance the robustness of the zero-watermarking scheme.

Advanced image processing attacks, such as blurring, median filtering, Gaussian filtering, denoising, and JPEG compression, were employed to simulate various image degradations and assess the algorithm's robustness under different conditions.

Furthermore, geometric attacks, including rotation, scaling, translation, and cropping, were performed to gauge the effectiveness of the zero-watermarking technique against spatial transformations and cropping scenarios.

The Bit Error Rate (BER) (9) serves as an evaluation metric to measure the detected bit errors between the original watermark (W) and the retrieved watermark (W'). A low BER value indicates a stronger level of watermark robustness [27]. In other words, a lower BER signifies that the retrieved watermark closely matches the original watermark, indicating a more reliable and accurate watermark retrieval process:

$$BER = \frac{Number\ of\ error\ bits}{Total\ bits}. \tag{9}$$

In addition, the Normalized Cross-Correlation (NCC) (10) is used to assess the similarity between the original watermark (*W*) and the retrieved watermark (*W'*) [28].The NCC measures how closely the extracted watermark matches the original watermark, providing a quantitative evaluation of their similarity:

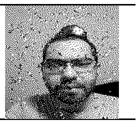
$$NCC = \frac{\sum(W(i,j)*W\prime(i,j))}{\sqrt{\sum(W(i,j)^2*\sum W\prime(i,j)^2)}}. \tag{10}$$

Fig. 9 showcases the Bit Error Rate (BER) and Normalized Cross-Correlation (NCC) results with different momentums on the batch normalization layers. After evaluating the neural network's performance, a momentum value of 4.5 was selected.

Table 2. Retrieved watermark advanced processing attacks.

Attack	Retrieved Watermark	Attack	Retrieved Watermark
JPEG Lossy Compression Quality Factor 90		Gaussian Filtering 3x3	
JPEG Lossy Compression Quality Factor 30		Median Filter 3x3	
Gaussian noise μ=0, σ^2=0.9		Blurring with re-scaling	

Table 3. Retrieved watermark geometric attacks

Attack	Retrieved Watermark	Attack	Retrieved Watermark
Rotation 0°-360°		Cropping by 150x150	
Circular shifting by *x*= 150		Centered cropping	
Down sampling 64 x 64		Translation x=100, y=100	
Translation x=100, y=0		Translation x=0, y=100	

This choice was based on its ability to yield a stable system with minimal variations in the BER and NCC values. Furthermore, this momentum value produced an average BER of 0.0046 and an average NCC of 0.9950, indicating highly accurate and reliable watermark retrieval.

The selection of momentum as 4.5 demonstrates its effectiveness in optimizing the network's performance, ensuring consistent and precise watermark recovery. The stable system achieved through this momentum value further increases the algorithm's robustness.

Table 4. Comparison applying geometric attacks

Parameter	[10]	[11]	[14]	[15]	Proposed
Detection metrics	BER, NC	BER, NC	NC	PSNR, MSE, NC	BER, NCC
Watermark size	32x32	16x16	32x32	64x64	160x160
Geometric attacks					
Cropping	Yes	Yes	Yes	Yes	Yes
Rotation	Yes	No	Yes	Yes	Yes
Scaling	Yes	Yes	Yes	Yes	Yes
Translation	No	Yes	Yes	Yes	Yes
Signal processing distortions					
JPEG	Yes	No	Yes	Yes	Yes
Gaussian noise	Yes	Yes	Yes	Yes	Yes
Median filter	Yes	Yes	No	Yes	Yes
Gaussian Filter	Yes	Yes	Yes	No	Yes
Blurring	No	No	No	No	Yes
Feature extraction method	Fractional Order-Legendre Fourier Moments	Statistical feature (skewness, entropy and median)	Neural networks	Neural networks	Neural networks
Medical images	No	Yes	Yes	Yes	Yes
Natural images	Yes	No	No	No	Yes

Fig. 10 and Fig. 11 provide a visualization of the obtained (BER) and (NCC) values from the testing using databases of medical imaging and natural images, respectively.

The BER values reflect the accuracy of watermark retrieval process for both medical and natural images. Fig. 11 displays the NCC values, demonstrating the similarity between the original watermark and the retrieved watermark.

The results demonstrate the algorithm's versatility in effectively handling different image formats, including NII, DICOM, and BMP respectively. The graphics exhibit the algorithm's remarkable robustness, reflected in the low error value and similarity between the original watermark and the retrieved watermark.

The Table 2 and Table 3 clearly demonstrate the efficient recovery of the watermark, regardless the image is tampered or distorted. This finding highlights the effectiveness and robustness of the proposed system for authenticating medical images. The watermark retrieval process obtained few errors, indicating that the system can accurately reconstruct the watermark even if the image has been distorted.

The efficiency of the watermark retrieval confirms the reliability and practicality of the proposed method for ensuring the authenticity and integrity of medical images. These results underscore the algorithm's ability to perform consistently and accurately across diverse types of images. Table 4 provides a comprehensive comparison of the robustness of the proposed algorithm against other methods.

The comparison with Hosny's [10] method focuses on a zero-watermarking algorithm for medical images using ROI characteristics. Additionally, the comparisons with Huang [15], Tayachi [11], and Gong [14] methods are emphasized since these schemes concentrate on obtaining characteristics from a pretrained neural models for master share generation.

The proposed scheme outperforms these methods in terms of robustness, even when subjected to various attacks. This demonstrates the algorithm's superior ability to withstand image alterations and maintain watermark integrity. The main advantage of the proposed system lies in the rigorous analysis of its performance through the modification of hyperparameters in the neural model.

This approach allows the system to focus on learning specific image characteristics, enabling better watermark recovery. The main disadvantage of this method lies in the training stage of the neural network for medical images.

The proposed scheme effectively generates protection for a set of medical images, it does not provide specific individualized protection for each image, since many of the features from the image can be obtain in other imaging analysis, this applies for most of the zero-watermarking schemes for medical imaging.

However, for natural images, the method may indeed offer individualized protection for every single image. The training process could potentially focus on capturing specific features and patterns unique to individual natural images, enabling more personalized image protection and authentication.

4 Conclusions

In this paper, a robust Zero-watermarking algorithm based on the Context Encoder neural network model is proposed. The algorithm obtains image features to enhance image reconstruction and increase the robustness of the zero-watermarking system.

We conducted extensive experiments encompassing medical imaging and natural images, including Neuroimaging Informatics Technology Initiative (NII) and Digital Imaging and Communications in Medicine (DICOM) formats for medical images and .bmp for natural images.

The results demonstrate a high level of robustness in both image types, with low Bit Error Rate (BER) values indicating reliable watermark retrieval. Additionally, the high Normalized Cross-Correlation (NCC) values signify a strong similarity between the original and retrieved watermarks, validating the algorithm's effectiveness in preserving watermark integrity during the retrieval process. Overall, the proposed algorithm exhibits robustness, efficiency, and versatility, making it a reliable solution for image authentication and verification.

With the neural model fine-tuning and the extraction of specific image features, the algorithm increased its efficiency and improved the watermark retrieval process, highlighting its potential for robust image protection and authentication. The results presented here validate the effectiveness of the proposed algorithm and its suitability for medical applications.

Finally, one of the key contributions of this research is the watermark with a size of 160 x 160 pixels, which surpasses the typical size employed in other research papers. The watermark size is particularly valuable as it enables medical professionals to easily identify the patient and authenticate the medical imaging.

The increased size provides a clear representation of the patient's face, facilitating quick and accurate verification by the doctor.

Acknowledgments

The authors thank the Instituto Politécnico Nacional (IPN), as well as the Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) for the support provided during the realization of this research.

References

1. **Neymeen, H., Boles, W., Boyd, C. (2013).** A review of medical image watermarking requirements for technology. Journal Digit Imaging, Vol. 26, No. 2, pp. 326–343.
2. **Coatrieux, G., Quantin, C., Montagner, J., Fassa, M., Allaert, F. A., Roux, C. (2008).** Watermarking medical images with

anonymous patient identification to verify authenticity. MIE, Vol. 136, pp. 667–672.

3. **Cedillo-Hernandez, M., Cedillo-Hernandez, A., Nakano-Miyatakea, M., Perez-Meana, H. (2020).** Improving the management of medical imaging by using robust andsecure dual watermarking. Biomedical Signal Processing and Control, Vol. 56. DOI: 10.1016/j.bspc.2019.101695.
4. **Garcia-Nonoal, Z., Mata-Mendoza, D., Cedillo-Hernandez, M., Nakano-Miyatake, M. (2023).** Secure management of retinal imaging based on deep learning, zero-watermarking and reversible data hiding. The Visual Computer, Vol. 40, pp. 1–16. DOI: 10.1007/s00371-023-02778-1.
5. **Sinhal, R., Ansri, I. A., Ahn, C. W. (2020).** Blind image watermarking for localization and restoration of color images. IEEE Access, Vol. 8, pp. 200157–200169. DOI: 10.1109/ACCESS.2020.3035428.
6. **Wang, H., Yuan, Z., Chen, S., Su, Q. (2023).** Embedding color watermark image to color host image based on 2D-DCT. Optik, Vol. 274. DOI: 10.1016/j.ijleo.2023.170585.
7. **Patel, A. K., Patel, P. (2022).** Color adaptive robust DWT-SVD watermarking algorithm and limitations: Color space comparisons. 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT), pp. 656–661.
8. **Sinhal, R., Ansari, I. A. (2023**). Machine learning based multipurpose medical image watermarking. Neural Computing and Applications, pp. 1–22. DOI: 10.1007/s00521-023-08457-5.
9. **Qasim, A. F., Meziane, F., Aspin, R. (2018).** A reversible and imperceptible watermarking scheme for MR images authentication. IEEE, 24th International Conference on Automation and Computing, pp. 1–6. DOI: 10.23919/IConAC.2018.8749000.
10. **Hosny, K. M., Darwish, M. M., Fouda, M. M. (2021).** New color image zero-watermarking using orthogonal multi-channel fractional-order legendre-fourier moments. IEEE Access, Vol. 9, pp. 91209–91219. DOI: 10.1109/ACCESS.2021.3091614.
11. **Tayachi, M., Nana, L., Pascu, A. C., Benzarti, F. (2023).** A hybrid watermarking approach for DICOM images security. Applied Sciences, Vol. 13, No. 10, pp. 6132. DOI: 10.3390/app13106132.
12. **Roček, A., Javorník, M., Slavíček, K., Dostál, O. (2017).** Reversible watermarking in medical imaging with zero distortion in ROI. 2017 24th IEEE International Conference on Electronics, Circuits and Systems, pp. 356–359. DOI: 10.1109/ICECS.2017.8292071.
13. **Fierro-Radilla, A., Nakano-Miyatake, M., Cedillo-Hernandez, M., Cleofas-Sanchez, L., Perez-Meana, H. (2019).** A robust image zero-watermarking using convolutional neural networks. 2019 7th International Workshop on Biometrics and Forensics (IWBF) IEEE. pp. 1–5. DOI: 10.1109/IWBF.2019.8739245.
14. **Gong, C., Liu, J., Gong, M., Li, J., Bhatti, U. A., Jixin, M. (2022).** Robust medical zero-watermarking algorithm based on residual-densenet. IET Biometrics, Vol. 11, No. 6, pp. 547–556. DOI: 10.1049/bme2.12100.
15. **Huang, T., Xu, J., Tu, S., Han, B. (2023).** Robust zero-watermarking scheme based on dephtwise overparameterized VGG network in healthcare information security. Biomedical Signal Processing and Control, Vol. 81.
16. **Han, B., Du, J., Jia, Y., Zhu, H. (2021).** Zero-watermarking algorithm for medical image based on VGG19 deep convolution neural network. Journal of Healthcare Engineering, pp. 104478. DOI: 10.1155/2021/5551520.
17. **Ganguly, K. (2017).** Learning generative adversarial networks: next generation deep learning simplified. Birmingham, United Kingdom, Packt Publishers Ltd.
18. **Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Ward-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2014).** Generative adversarial networks. Computer and Information Sciences.
19. **Arjovsky, M., Chintala, S., Bottou, L. (2017).** Wasserstein generative adversarial networks. Proceedings of the 34th International Conference on Machine Learning, PMLR, Vol. 70, pp. 214–223.

20. **Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A. A. (2016).** Context encoders: Feature learning by inpainting. Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2536–2544.
21. **Dohi, N., Rathnayake, N., Hoshino, Y. (2022).** A comparative study for COVID-19 cases forecasting with loss function as AIC and MSE in RNN family and ARIMA. 2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems (SCIS&ISIS), IEEE, pp. 1–5. DOI: 10.1109/SCISISIS55246.2022.10001870.
22. **Lo, S. Y., Patel, V. M. (2021).** Error diffusion halftoning against adversarial examples. arXiv. pp. 3892–3896. DOI: 10.1109/ICIP42928.2021.9506591.
23. **Arevalo-Ancona, R. E., Cedillo-Hernandez, M. (2023).** Zero-watermarking for medical images based on regions of interest detection using K-means clustering and discrete fourier transform. International Journal of Advanced Computer Science and Applications, Vol. 14, No. 6. DOI: 10.14569/IJACSA.2023. 0140662.
24. **Medical Segmentation Decathlon**. http://medicaldecathlon.com/.
25. **SIIM-ACR Pneumothorax Segmentation**. https://www.kaggle.com/competitions/siim-acr-pneumothorax-segmentation/data.
26. **FREE STOCK IMAGES**. https://www.stockvault.net.
27. **Dai, Z., Lian, C., He, Z., Jiang, H., Wang, Y. (2022).** A novel hybrid reversible-zero watermarking scheme to protect medical images. IEEE Access, Vol. 10, pp. 58005–58016. DOI: 10.1109/ACCESS.2022.3170030.
28. **Morales-Ortega, A., Cedillo-Hernandez, M. (2022).** Ownership authentication and tamper detection in digital images via zero-watermarking. 2022 45th International Conference on Telecommunications and Signal Processing IEEE, (TSP), pp. 122–125. DOI: 10.1109/TSP55681.2022.9851253.
29. **Oueslati, S., Cherif, A., Solaimane, B. (2011).** Adaptive image watermarking scheme based on neural network. International Journal of Engineering Science and Technology, Vol. 3, No. 1, pp. 748–757.
30. **Hosny, K., Darwish, M. (2021).** New geometrically invariant multiple zero-watermarking algorithm for color medical images. Biomedical Signal Processing and Control, Vol. 70. DOI: 10.1016/j.bspc.2021.103007.

Article received on 04/07/2023; accepted on 13/10/2023.
[*]Corresponding author is Manuel Cedillo-Hernandez.