Zero-Watermarking for Medical Images Based on Regions of Interest Detection using K-Means Clustering and Discrete Fourier Transform

Rodrigo Eduardo Arevalo-Ancona, Manuel Cedillo-Hernandez
SEPI-ESIME Culhuacan, Instituto Politecnico Nacional, Mexico city, Mexico

Abstract—Watermarking schemes ensure digital image security and copyright protection to prevent unauthorized distribution. Zero-watermarking methods do not modify the image. This characteristic is a requirement in some tasks that need image integrity, such as medical images. Zero-watermarking methods obtain specific features for the master share construction to protect the digital image. This paper proposed a zero-watermarking scheme based on K-means clustering for ROI detection to obtain specific features. The K-means algorithm classifies the data according to the proximity of the generated clusters. K-means clustering is applied for image segmentation to identify ROI and detect areas that contain important information from the image. Therefore, the Discrete Fourier Transform (DFT) is applied to the ROI features, using the high frequencies to increase its robustness against geometric attacks. In addition, an edge detection based on the Sobel operator is applied for the QR code creation. This type of watermark avoids errors in watermark detection and increases the robustness of the watermark system. The master share creation is based on an XOR logic operation between extracted features from the selected ROI and the watermark. This method focuses on the protection of the image despite it being tampered with. Many proposed schemes focus on protection against advanced image processing attacks. The experiments demonstrate that the presented algorithm is robust against geometric and advanced signal-processing attacks. The DFT coefficients from the extracted ROI features increase the efficiency and robustness.

Keywords—Zero-watermark; ROI detection; machine learning; k-means; image security; copyright protection

I. INTRODUCTION

In recent years, the advances in communication technologies and multimedia file sharing through different digital systems have increased related to the conditions generated by the COVID-19 pandemic. In addition, the images transmitted through different communication channels may contain sensitive information [1]. For this reason, technology for digital image protection, authentication and copyright protection is a requirement. One technology that has attracted the attention of researchers is the watermarking systems, which provide security, copyright protection, or certify digital images [2].

The protection of medical images has become a relevant task in recent years since there has been an increase in remote medical consultations. Therefore, medical images carried out the patient's data which is an essential requirement. For this reason, watermarking systems are a solution to this problem.

Traditional embedding algorithms imperceptibly embed ownership information into the host image to ensure the copyright, consequently, the signal is recovered from the watermarked image [3], [4], [5]. Traditional watermarking methods embed a signal with ownership information into a host image [4], [5]. This process distorted the image and modified its information. However, this process can generate some distortion generating a wrong image analysis. An example is the one proposed by Juarez-Sandoval et al. [3]. They present a method of imperceptible-visible watermarking. A homogeneous region is detected by the variance of the values in the pixels, followed by the just noticeable difference (JND), which represents the maximum luminance variation. The JND identifies the most appropriate area in the image to embed the binary watermark. Guanghui and Hao [6] used the Arnold transform to encrypt the watermark. On the other hand, they applied the Wavelet Transform based on the Mallat decomposition, representing the coefficients of the low-pass and high-pass filters, thus adjusting the intensity of the watermark to the pixel variations. To avoid distortions in the images, Zero-watermarking techniques are developed, these schemes do not embed information into the digital data. Instead, zero-watermarking used specific features from the image, and the watermark to create a master share (a feature matrix) without losing the host image quality. The master share is unique for each image.

Region of interest (ROI) detection on images identifies specific areas with relevant information for its analysis and features extraction that belongs to the selected region [7], [8]. On the other hand, regions of non-interest (RONI) generally used related features to the background [9], [10]. ROI methods are used in watermarking algorithms to make the embedding and detection process of the watermark signal more efficient since it takes advantage of the detected features that are unique for each ROI. Zhang et al. [11] developed a watermarking system by inserting the signal into the RONI of medical images. They applied the Discrete Wavelet Transform (DWT) and added two bits in each frequency sub-band obtained (low, medium, high). Next, the Otsu algorithm identifies the RONI, and then the detected ROI is encrypted with a hashing algorithm to combine it with the patient's information. In [12] Qi et. al. proposed two factors to select the ROI by detecting variations, especially in the case of medium frequencies. Subsequently, the visual effect factor (VEF) determines the
region of watermark embedding. Lampezhiev et al. [13], proposed a K-means segmentation for ROI detection on medical diagnosis. Therefore, a fuzzy clustering evaluation criterion was applied to select specific features from the image to determine the statistical data required for making decisions in applied medicine.

Medical image watermarking schemes divide the image into ROI and RONI for signal embedding. In addition, the ROI and RONI segmentation can be modified easily by the software. As a solution, zero-watermarking systems generate lossless protection in the image quality being more efficient. In zero-watermarking algorithms, the signal is not embedded into the base image [14], [15]. Extracted features from the image are fuzzed with the watermark related to the owner’s information for the creation of the master share. These associations are stored and provide continuous protection [16]. In addition, the main advantage of this algorithm is the generated robustness.

Khafaga et al. presented in [17] a descriptor based on multi-channel Gaussian-Hermite moments of fractional order for feature extraction to create a vector with the most robust features and used the 1D Chebyshev chaotic map to scramble the watermark and increase its security. Finally, it is performed an XOR operation for the master share creation. Xing, Li, and Liang created a zero-watermarking scheme [18]. The Discrete Cosine Transform (DCT) is applied to high-frequency coefficients obtained from the Discrete Fourier Transform (DFT). The coefficient matrix is extracted from the left corner and has the same size as the watermark. The Arnold Transform scrambles the watermark to increase the security of the system. Thus, the coefficient matrix and the scrambled watermark are fuzzed for the master sharer generation.

ROI detection-based algorithms have a relevant role in zero-watermarking systems, as it identifies unique features from each image [15]. The extracted features create the master share serving for identification, protection, authentication, and certification of the digital image against misuse [19], [20]. Fang et al. in [21] present a watermarking scheme that detects specific areas in medical images by extracting the SIFT descriptors. Therefore, they apply the Bandelet Transform for pixel change detection. Thus, the Discrete Cosine Transform (DCT) is applied to generate more robustness against geometric attacks. The Arnold Transform increases watermark security. Finally, the watermark and the ROIs are combined. Gong et al. [22] applied a Residual-DenseNet to obtain a feature vector of the image, then the logistic map creates a Chaotic matrix and generates the feature matrix through a logical XOR operation, which is stored to verify the image. In [23] Hosny and Darwish applied Multi-channel Fractional-order Gegenbauer moments of color images to obtain the ROI-related features of the image and form the feature vector, which is combined with the watermark to generate image protection.

This paper proposes a zero-watermarking scheme-based detection of regions of interest using K-means clustering. Therefore, the DFT is used to obtain high frequencies to create a feature matrix. The features remain without distortions if the image has been tampered with. Thus, the process of the watermark construction used the Sobel filter to obtain the image edges. Finally, the master share is created by fuzzing the watermark and the feature matrix.

The k-means algorithm provides robustness to the presented method, generating an image segmentation based on features clustering for the ROI detection. The main advantage of this technique is the ROI detection which remains without modifying the pixel values used for the master share generation and detection even if the image has been tampered with. On the other hand, the DFT makes the system more efficient, even though some geometric or advanced image processing attacks are applied to the image by an unauthorized user. High frequencies do not change despite image manipulation. The master share provides continuous copyright protection and image certification.

The main contributions of this paper are:

- K-means ROI detection and DFT-based feature extraction increase the robustness of the watermark system.
- ROI detection creates a feature matrix related to the image. These features are not modified when advanced signal processing is applied to the image.
- K-means clustering is applied for image segmentation to identify ROI and image patterns for the identification of areas with important infrastructures from the image.
- The DFT coefficients generate unique invariant features against geometric attacks.
- The extracted features for the master share construction are unique and generate a lossless watermarking system, which does not distort or modify the image, allowing its analysis for a correct diagnosis.
- This method focuses on the protection of the image despite it being tampered with. Many proposed schemes focus on protection against advanced image processing attacks.

The rest of the paper is organized as follows. Section II provides the background of study. Section III presents the proposed method. Section IV provides the experimental results, and Section V concludes this paper.

II. BACKGROUND

This paper proposed a zero-watermarking scheme based on K-means clustering for ROI detection. ROI detection increases the efficiency of feature detection and extraction. Furthermore, high-frequency coefficients of the DFT provide robustness against geometric attacks, where the selected features do not change.

A. K-Means Clustering for Image Segmentation

Image segmentation analyzes the image and identifies ROI with useful detected features [24]. ROI detection may determine the areas which must be tampered avoiding their use for feature extraction to create the master share [25]. ROIs have areas with relevant information. In the case of medical images, this information is vital for analysis for a correct and
efficient diagnosis, so they should not be modified or distorted and thus not affect the patient’s diagnosis.

Unsupervised learning refers to a kind of machine learning where there are no labels or the output is not known, like the clustering algorithms [26], this technique is like classification methods since it divides the data into groups called clusters.

Unsupervised learning refers to a kind of machine learning where there are no labels or the output is not known, like the clustering algorithms [26], this technique is like classification methods since it divides the data into groups called clusters.

Clustering techniques separate the data by associating points into different classes. K-means is one of the most used clustering algorithms, due to its implementations which consist of an iterative algorithm that assigns the data to a specific cluster based on the Euclidean distance (1) from an arbitrary centroid ($\eta$).

$$\min(E(\eta)) = \sqrt{\sum_{i=1}^{n}(p_i - \eta)^2}$$  \hspace{1cm} (1)

where $p$ is the data point, and subsequently the centroids are realigned to the mean of the assigned clusters ($S$) on each iteration (2) [27], [28].

$$\frac{\delta E}{\delta \eta} = 0 \rightarrow \eta_{i+1} = \frac{1}{S} \sum_{p}$$ \hspace{1cm} (2)

K-means is a popular algorithm because is easy to understand and implement and it can be used in many tasks. One of the disadvantages is that the number of centroids must be set before the initialization. K-means algorithm clusters the pixels for image segmentation and pattern recognition for ROI detection [29], [30]. The clustering technique detects interest points on the image for pattern recognition using different centroids [31]. K-means segmentation detects ROI to determine the most important features from the image to obtain better results for pattern recognition and feature extraction in zero-watermarking algorithms [32] (Fig. 1). Segmentation methods recognize ROI if an attack is applied to the image the feature points may change. Furthermore, DFT coefficients improve feature detection due to their properties which make the frequency coefficients from the selected area invariant to geometric attacks.

### B. 2D Discrete Fourier Transform

The Discrete Fourier Transform (DFT) (3) makes a representation of the space domain image into the frequency domain, providing robustness against geometric attacks (scaling, cropping, translation, rotation) [33].

$$F(u, v) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) e^{-j2\pi \frac{ux}{M} \frac{vy}{N}}$$ \hspace{1cm} (3)

The Fourier Spectrum (4) (Fig. 2) describes frequency coefficients. The DFT has a real value (Re) and an imaginary value (Im).

$$|F(u, v)| = \sqrt{(\text{Re}(F(u, v)))^2 + (\text{Im}(F(u, v)))^2}$$ \hspace{1cm} (4)

Phase (5) describes the symmetry of the signal.

$$\theta = \tan^{-1}\left(\frac{\text{Im}(F(u, v))}{\text{Re}(F(u, v))}\right)$$ \hspace{1cm} (5)

The magnitudes in (4) and (5) make a polar representation of the DFT (6).

$$F(u, v) = |F(u, v)| e^{-j\theta uv}$$ \hspace{1cm} (6)

The DFT has the properties of linearity, scale, translation, symmetry, rotation, and cropping (sampling). Consequently, the DFT improves the performance and efficiency of the presented method.

### III. PROPOSED ZERO-WATERMARKING SCHEME

This paper presents a zero–watermarking scheme, focusing on the master share generation aimed to protect digital images. K-means algorithm detects image ROI for feature extraction (Fig. 3).

The selected features from the ROI are robust against image processing attacks. The DFT makes a robust zero-watermarking algorithm, and high-frequencies coefficients are used for master share creation.
C. Region of Interest Feature Extraction

The master share generation (Fig. 4) is based on K-means ROI detection which provides robustness against advanced image processing attacks since the selected features do not change. In addition, combining ROI detection and DFT increases the robustness of the watermark system. Consequently, this method improves the performance and efficiency of the presented method.

The idea of the K-means implementation is to detect the main ROI from the medical image to extract the main image features. To improve the robustness of the watermark the ROI are eliminated. Therefore, the DFT is applied to the new image matrix to obtain the high-frequency coefficients and increase the robustness against geometric attacks. DFT domain correlated patterns to improve the image details. In addition, the matrix with the DFT coefficients is divided into non-overlapping 8x8 blocks for its binarization for the master share construction, using the mean value as a threshold (Fig. 4). In addition, the DFT coefficients are sorted from the higher to lower frequencies, and a matrix with a size of 160 x 160 is created.

D. Watermark Construction

The watermark is unique for each image using the characteristics from the image. Generating this type of watermark avoids conflicts in the detection stage and increases the system's robustness. Therefore, it becomes a useful method for image security and copyright protection. The watermark generation process consists of the patient's information encrypted in a QR code with an image of the edge detection with the Sobel filter on the QR code center (Fig. 5).

QR codes are modules in which the patient's information can be stored, as well as their doctor, or redirect the user to their electronic file. On the other hand, QR codes can be detected by different devices, such as cell phones, computers, or tablets regardless of a loss of information.

In addition, these types of identification codes are quickly accessible, facilitating the identification of patients. Therefore, the QR codes can be used as a watermark for images.

The features from the edge detection reduce the watermark retrieval error and avoid ambiguity since the watermark is unique for each image. Furthermore, edge detection identifies key points from the image and applies them for image protection. In addition, the watermark is resized to a matrix of 52 x 52. These features increased the efficiency of the zero-watermarking scheme.

E. Master Share Generation

The master share is the element that provides the image security, certifies it, and protects the copyright protection, and must be stored in an external device. The master share (MS) is generated by fuzzing the image's unique features (imf) with the constructed watermark (Ws) with an XOR (⊕) logic operation (7). The image features are stables and invariants as a requirement.

\[ MS = imf \oplus Ws \] (7)

The master share (Fig. 6) is unique for each image and is created for the efficient protection of the digital image.

F. Master Share Detection

In the detection phase (Fig. 7.), the ownership authentication is validated verifying the image and certifying its authenticity. The generated master share and the unique features reveal the watermark. The watermark is detected by applying the logical XOR operation between the image features and the stored master share corresponding to the original image.

The watermark is recovered using the extracted features and the Master Share, this procedure is described as follows.

Fig. 6. A) Constructed watermark, B) Master share.
Step 1: The image is segmented using the K-means clustering algorithm for the localization of the regions of interest.

Step 2: The DFT high-frequency coefficients are obtained.

Step 3: The extracted features are binarized.

Step 4: It is applied an XOR operation between the binarized sequence and the Master Share.

Step 5: The watermark is recovered, and it is verified.

IV. EXPERIMENTAL RESULTS

Many experiments were realized to evaluate the performance of the proposed algorithm. The testing image database contains 708 images with a size of 512 X 512. In addition, the watermark is 160 x 160. The dataset was obtained from [34].

Some advanced image processing (blurring, median filter, gaussian filter, denoising, jpeg compression) and geometric attacks (rotation, scale, translation, cropping) were performed to evaluate the zero-watermarking scheme.

As an evaluation metric the bit error rate (BER) (8) is used to measure the detected bit errors between the watermark (W) and the retrieved watermark (W'). A low BER indicates stronger watermark robustness [16].

{\begin{equation}
BER = \frac{\text{Error bits}}{M \times N}
\end{equation}}

(8)

In addition, the normalized cross-correlation (NCC) (9) evaluates the similarity between the detected watermark and the extracted watermark [19].

{\begin{equation}
NCC = \frac{\sum_{i,j} W(i,j) W'(i,j)}{M \times N}
\end{equation}}

(9)

where m and n are the dimensions of the watermark.

Table I and Table II demonstrate the robustness of the proposed zero-watermarking scheme. The BER and NCC metrics ensure the effectiveness of the algorithm.

The BER value is closer to 0, indicating that the recovered watermark has a low error. The NCC is closer to 1, showing a great similarity between the detected watermark and the embedded watermark.

TABLE III. RETRIEVED WATERMARK AGAINST DIFFERENT TAMPERING ATTACKS

<table>
<thead>
<tr>
<th>Tampere d Image</th>
<th>Watermark</th>
<th>Retrieved Watermark</th>
</tr>
</thead>
<tbody>
<tr>
<td>No attack</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roll translation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median filter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale 170 x 170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian noise</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table II. ADVANCED IMAGE PROCESSING ATTACKS ROBUSTNESS TEST

<table>
<thead>
<tr>
<th>Attack</th>
<th>BER</th>
<th>NCC</th>
<th>Attack</th>
<th>BER</th>
<th>NCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPEG 90</td>
<td>0.0041</td>
<td>0.9958</td>
<td>Scale and Blurring</td>
<td>0.0094</td>
<td>0.9900</td>
</tr>
<tr>
<td>JPEG 70</td>
<td>0.0053</td>
<td>0.9950</td>
<td>Gaussian filter</td>
<td>0.0041</td>
<td>0.9955</td>
</tr>
<tr>
<td>JPEG 30</td>
<td>0.0061</td>
<td>0.9941</td>
<td>Scale and Gaussian filter</td>
<td>0.0071</td>
<td>0.9929</td>
</tr>
<tr>
<td>Gaussian noise</td>
<td>0.0060</td>
<td>0.9930</td>
<td>Denoising</td>
<td>0.0062</td>
<td>0.9939</td>
</tr>
<tr>
<td>Scale and Gaussian noise</td>
<td>0.0054</td>
<td>0.9945</td>
<td>Median filter</td>
<td>0.0052</td>
<td>0.9946</td>
</tr>
<tr>
<td>Blurring</td>
<td>0.0070</td>
<td>0.9923</td>
<td>Scale and median filter</td>
<td>0.0056</td>
<td>0.9947</td>
</tr>
</tbody>
</table>

TABLE I. GEOMETRIC ATTACKS ROBUSTNESS TEST

<table>
<thead>
<tr>
<th>Attack</th>
<th>BER</th>
<th>NCC</th>
<th>Attack</th>
<th>BER</th>
<th>NCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No attack</td>
<td>0.0058</td>
<td>0.9944</td>
<td>Bottom left crop</td>
<td>0.0077</td>
<td>0.9929</td>
</tr>
<tr>
<td>Radom rotation</td>
<td>0.0091</td>
<td>0.9900</td>
<td>Upper right crop</td>
<td>0.0055</td>
<td>0.9941</td>
</tr>
<tr>
<td>Roll translation 150</td>
<td>0.0064</td>
<td>0.9942</td>
<td>Center crop</td>
<td>0.0072</td>
<td>0.9931</td>
</tr>
<tr>
<td>Scale</td>
<td>0.0053</td>
<td>0.9945</td>
<td>Translation</td>
<td>0.0095</td>
<td>0.9904</td>
</tr>
</tbody>
</table>

www.ijacsa.thesai.org
Table III shows the efficiency of the watermark detection stage and the robustness of the proposed method. The retrieved watermark is very similar to the original watermark, which can be seen in Table I and Table II. The efficiency of the zero-watermarking scheme is similar regardless of the selected clusters.

Fig. 8 shows the variations on the BER with different numbers of clusters against different advanced image processing attacks and geometric attacks. Fig. 9 demonstrates how the geometric certification stage recovers a similar retrieved watermark to the original watermark. Hence the selection of clusters will depend on the task and the image type. Since the selection of ROI will depend on the requirements of the application of the watermarking system. ROI selection may be necessary to determine a zone with user-specified characteristics for master share creation, as in Fig. 10.

The proposed algorithm is compared with other ROI zero-watermarking lossless algorithms. The comparison of our proposed zero-watermarking scheme with other schemes (Zhang et al. [35], Huang et al. [36], Cheng et al. [37], Zhou et al. [38] and Jing et al. [39]) was made in terms of the following aspects: 1) robustness against geometric attacks (no attack, rotation, translation, scale, crop). 2) Robustness against image processing attacks (blurring, noise addition, and JPEG 30 compression). 3) The BER is used for the analysis of the recovered watermark. 4) To evaluate its similarity with the original watermark, the NCC was used. 5) For a fair comparison, the same dataset is used and the same conditions from the experimental environment. In addition, some of the authors from the comparison schemes increased the robustness of their algorithms, this can be observed in the variations from the BER values and the similarity measures that they presented. The results are in Fig. 11 and Fig. 12.

The proposed method presents a better performance against different geometric and advanced image processing attacks.

The high-frequency coefficients from the DFT provide robustness against geometric attacks and the ROI detection increases the efficiency against advanced image processing attacks.
The method presented in this paper better satisfies the lossless requirements and provides robustness against the different geometric and advanced image processing attacks (Table I and Table II). Comparing our method with other zero-watermarking has a better performance. The recovered watermark can be easily distinguished and provides security for copyright protection related to digital images. On the other hand, the features related to the image as a watermark provide an effective protection system.

V. CONCLUSIONS

In this paper, a zero-watermarking algorithm is presented based on ROI detection for image certification and authentication. The results demonstrate the robustness of the zero-watermarking system and the similarity of the retrieved watermark, generating continuous image protection, verification, authentication, and certification. On the other hand, the features related to the image as a watermark increases image security. Moreover, the ROI detection based on K-means generated a better performance of the zero-watermarking system. The results show a minimum loss on the watermark recovery. The use of different ROI areas does not significantly modify the results obtained, however, when using a greater number of clusters, the processing time increases.

In future work, the use of a specific ROI for the extraction of the characteristics is proposed. The user can focus on image analysis to perform artificial intelligence tasks related to robotics and computer vision. The system would generate a verification method for images related to security areas and avoid their misuse. In the same way, image databases could be generated for various tasks with protection and a system for user verification.

ACKNOWLEDGMENT

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

REFERENCES


