International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Jul-24)

http://dx.doi.org/10.12785/ijcds/160140

A Secure Self-Embedding Technique for Manipulation Detection and Correction of Medical Images

Afaf Tareef[1]

[1]*Faculty of Information Technology, Mutah University, Jordan*

Received 20 Feb. 2024, Revised 21 Apr. 2024, Accepted 28 Apr. 2024, Published 1 Jul. 2024

Abstract: The protection of medical images transmitted through the E-healthcare system is very critical. Nowadays, medical image watermarking has been emerged as trustworthy way to authenticate medical information during transmission. This paper presents a secure self-embedding scheme that detects and corrects the tampesr in medical images. The proposed scheme involved two decomposition and dimensionality reduction techniques, singular value decomposition and learning sparse decomposition. First, the color medical image is transformed into YCrCb color space and the luminance plane is chosen. To create the watermark, the medical image is automatically classified into region of interest (ROI) and region of non-interest (RONI), and then, the ROI is encoded by sparse decomposition with convolutional Basis Pursuit DeNoising (BPDN) dictionary. The sparse watermark is then hidden in the singular values of the host part of the image. The quantitative and qualitative results show that the proposed method is robust against numerous aggressive and geometric distortions without compromising the quality of the original medical image. The proposed algorithm yields a high Peak Signal-to-Noise Ratio (PSNR) larger than 45dB for all type of images, as well as high normalized correlation (NC) value under all types of attacks. It is demonstrated that the presented system performs better than the existing techniques, and could be helpful for e-healthcare systems.

Keywords: Color medical images, Automatic Segmentation, Self-Embedding, Manipulation Detection and Correction, Encryption.

1. INTRODUCTION

Telemedicine is a rapidly quickly emerging field where medical images is transmitted through the electronic healthcare environment for the intention of diagnosis, consulting, and remote medical intervention. In this context, the secure transmission of medical images has a valuable influence on the proper examination of serious diseases besides reducing the misdiagnosis problems. Nowadays, medical images watermarking has become a reliable method of verifying medical data during transmission. Numerous watermarking systems for medical images have been introduced based on efficient mathematical transformations, including Fourier transform, cosine transform, wavelet transform, singular value decomposition, compressed sensing, and sparse decomposition.

Singular Value Decomposition (SVD) is among the most effective methods available for protecting sensitive data. SVD is a factorization method used for dimensionality reduction by extracting algebraic feature from images [1]. It factorizes a 2D matrix M into three identically sized matrices, one diagonal matrix and two unitary matrices, as shown in Eq (1):

$$M = U \Lambda V^T = \sum_{i=1}^{r} \lambda_i u_i v_i^T \tag{1}$$

where Λ is a singular value matrix of positive real numbers, whereas the other two matrices (i.e., U and V) are the eigenvectors. Singular value decomposition has been extensively utilized in watermarking generally, and medical image protection specifically, thanks to the stability characteristic of the singular values and the high resistance to geometric attacks.

Another effective and reliable tool recently gain a great attention is the sparse decomposition (SD). SD with all its versions has been unfolded as a promising transform utilized to effectively present and compress high-dimensional signals. Generally, SD factorizes a given vector into a linear combination of a few sparse vectors from an over-complete dictionary [2]. Specifically, each vector $\overrightarrow{\xi} \in R^K$ can be basically computed by the following formula:

$$\overrightarrow{\xi} \cong \sum_{i=1}^{n} \overrightarrow{b_i}\, s \tag{2}$$

E-mail: a.tareef@mutah.edu.jo

where $\overrightarrow{b_i} \in R^K$ is the i^{th} dictionary vector and $\overrightarrow{s} \in R^n$ is the sparse vector. As the dictionary is overcomplete, i.e.,($n > k$), it can identify a wide range of patterns in the input data. Over the past few decades, SD has established its value and effectiveness in a variety of signal processing fields, such as compression, Denoising, pattern classification and blind source separation. In 2014, a new era has begun in image watermarking field by the innovative method presented in [3], where sparse coding was utilized in order to compress and encrypt the watermark before hiding in the host image.

This paper introduces a tamper detection and correction scheme based on SVD and SD transformations. Sparse decomposition with convolutional learned dictionary is utilized to encrypt and compress the automatic-segmented ROI before hiding in the medical image. This leads to improve the embedding capacity, and achieve a strong encryption of the ROI. It also reduces the perceptual degradation in the watermarked image. This is accomplished by minimizing the number of elements required to represent the hidden information. Moreover, because noise cannot be sparsely represented, applying SD in watermarking helps in achieving the necessary robustness against noise.

2. Literature Review

Protecting medical images against manipulation in e-healthcare systems is crucial need to prevent incorrect diagnoses and treatments. Thus, various techniques have been presented to confirm the authenticity of medical data based on signature or hash [4], encryption [5], [6], and watermarking (e.g., [7], [8], [9]). For instance, Jabbar et al. [4] suggested a frequency domain method that combines the hash signature with the watermarking system. First, robustness hash code was generated based on Slice Transform (SLT) and Discrete Cosine Transform (DCT). Next, it was coupled with the patient record in order to form a watermark, which is then used to create a signature by encoding it based on a chaotic map with a secret key.

El-Shafai et. al. [5] proposes an optical-based authentication approach for medical images protection based on useful hashing, steganography, and encryption algorithms. This method employs sparse decomposition technique to compress the color elements of the image. Also, it performs the sigmoid quantization to the compressed medical image in order to create the digital quantized elements. To increase its secrecy, the message is then encrypted by Rubik's cube-based method to get the encrypted medical image. However, the rising frequency of security attacks in e-healthcare systems show that these methods lack efficiency.

The regions in the medical images can be classified into region-of-interest (ROI) and region-of-non-interest (RONI), where ROI plays the main role in diagnosis. In contrast, the RONI is either insignificant or not helpful for diagnosis and treatment purpose. Various ROI-based watermarking methods have been presented in the literature, which are performed in either spatial domain or frequency domain. Spatial domain-based fragile watermarking has been used for medical image authentication where the watermark has the ability to destroy itself in the event of any kind of attack, whether malicious or not. Depending on the embedding strategy, fragile watermarking can be classified into two types: block-based and pixel-based watermarking. Many block-based techniques have been proposed to insert a fragile watermark into the non-overlapping block of the image [10], [11], [12]. For instance, Gull et al., [12] presented another fragile watermarking approach to localize the tamper in the gray and color images. The medical image is first split into 4×4 blocks, and each block is further split into two blocks; upper block and lower block. The data for manipulation detection is hidden in the lower block, while the information for localization is hidden in the upper block. A block-based tamper detection results in a significant false positive rate because of identifying the entire block as a tampered if just one pixel is altered. This led to the development of the pixel-based fragile watermarking, where the watermark extracted from the original pixel is embedded into the pixel itself [13]. In [14], a blind fragile watermarking method for tamper detection is presented. The Speeded Up Robust Features (SURF) algorithm is applied to the ROI. Then, the binary message is hidden into the blocks around the SURF points. A semi-fragile watermarking method for JPEG2000 image is proposed in [15]. In this method, a perceptual hash function (PHF) is utilized on the wavelet elements and used to generate the watermark, which is then concealed in the host image during the JPEG2000 compression process. In general, fragile watermarking techniques achieve a high imperceptibility, however, they cannot withstand different kind of attacks. Moreover, the fragile and semi-fragile watermarking methods are not able to correct and reconstruct the tampered area in the attacked image.

Robust watermarking based on frequency domain has been successfully applied for tamper detection purpose. For instance, Parah et al. [7] introduced a ROI-based tamper detection using block-based cosine transform. The watermark is hidden in the chosen DCT coefficients of 8×8 blocks. A dual watermarking method using stationary wavelet transform and SVD is introduced in [8] for color general and medical images. This algorithm hides two watermarks: gray-scale watermark concealed in the green layer of the color image for the copyright protection purpose, and binary watermark concealed in the blue channel for tamper detection purpose. The genetic algorithm is also applied to achieve better perceptual quality and robustness. Swaraja et al. [9] suggested an algorithm for medical image authentication based on Discrete Wavelet Transform (DWT) and Schur transform. This method conceals dual watermarks into the blocks of the RONI. Particle Swarm Bacterial Foraging Optimization algorithm is also used to select the optimal threshold. To enhance the invisibility, Lempel-Ziv-Welch compression algorithm is performed. Another medical image watermarking method is proposed in [16] which incorporates two intelligent algorithms, which are genetic algorithm and particle swarm optimization to take advantage of the relationship between the pixel values of an image. However, the performance of such method mainly depend on chosen the optimal scaling factor used for embedding.

Another frequency domain-based watermarking system for medical image authentication is proposed in [17], where the wavelet transform and fractal dimension theory are used. The regularity of the mutation structure in the medical image is examined using a measure specified in the frequency domain. Additionally, the elements of the restored wavelet values are reduced using the fractal dimension. Based on the singularity of the fractal dimension of the block data, the main characteristics are determined and the fractal characteristic is created as the image authentication characteristic. Singh [18] introduces an invisible watermarking for color medical image based on DCT and lifting wavelet transform (LWT). First, the digital signature and patient report are combined and hidden into the host medical image for identity authentication purpose. In addition, the patient report is encrypted using BCH error correcting code and the signature watermark is encrypted using message-digest before concealing into the host image.

Various watermarking systems have been suggested in the literature for both tamper localization and correction. For instance, a block-based tamper recovery method is proposed in [19]. First the ROI hash is computed to find the tamper. The pixels of each ROI block are concealed into the least significant bits (LSBs) of the corresponding RONI block. The ROI hash value and the hash value of the reconstructed region are compared during the extraction process, and only modified blocks are recovered. The average and variance of each block are used to identify the modified blocks. Khor et al. [20] introduced a tamper detection and correction method for medical image, where the ROI information is embedded into RONI least significant bits. Another LSB-based watermarking method is proposed in [21], where the watermark is generated by perfuming exclusive-OR function between the mean pixel value of 4×4 pixel blocks and a key generated by the logistic map (LM). The watermarks bits are then hidden in LSBs of the image blocks. In general, watermark embedding in the spatial domain is less robust than embedding in the frequency domain, and the ROI cannot be correctly recovered if the RONI is tampered.

Zero-watermarking and Error Correcting Code are used in 2021 by [22] for tamper detection. First, RS code is directly applied on the ROI pixels to be stored in RONI. The parameters of RS code are guardedly selected in order to obtain a good restoration capability. The performance of this method including the recovery ability is based mainly on the RS code parameters. Also, this method is fragile against some attacks (e.g., rotation and JPEG compression). Another ROI-based tamper correction method for gray medical image is proposed in [23] using principal components (PC) for robust watermark and Lempel–Ziv–Welch method for fragile watermark. Moreover, a self-embedding watermarking system is presented in 2022 by [24] based on a spiral block mapping. In this method, a 3×3 block-based coding is performed to establish a watermark consisting of two authentication bits and seven restoration bits embedded in the LSBs. The spiral block mapping is used to hide the recovery bits into number of sub-blocks. This method is evaluated under several attacks such as blurred image, noise addition, and sharpening, and it yields high robustness and invisibility results. In [25], a reversible watermarking is introduced for tamper correction based on Scale-Invariant Feature Transform (SIFT). The image is first divided into blocks, and the features of each block are extracted to be concealed into the corresponding feature region. These methods achieve high imperceptibility level and low complexity time. However, they cannot restore the original tampered area. In addition,they have very low embedding capacity.

In 2023, many watermarking methods have been developed for accurate restoration of tampered ROI in the medical images. For instance, a zero-watermarking method is introduced based on K-means clustering [26]. The medical image is classified by K-means algorithm to identify ROI that contain important information. Then, Fourier transform is performed to the ROI to improve the robustness against geometric attacks. Next, Sobel operator is utilized to obtain the edges for the QR code creation. The embedding is performed by applying XOR operation between extracted ROI edges and the watermark. Another ROI-based tamper correction scheme is proposed in [27]. The ROI in the medical image is manually segmented then compressed and hidden in the RONI region. In addition, a small amount of data is hidden in ROI to identify ROI tampering. In general, the ROI-based methods show a high invisibility and a good capability of ROI recovery. However, most of the current methods in the literature depend on a manual selection of the ROI, which is error-prone process.

An accurate automatic segmentation of the ROI still remains a challenging problem. In this paper, the ROI is automatic segmented and post-processed by morphological operations. The ROI is then compressed and hided in a stable part of the host image for effective authentication and restoration of gray and color medical image. Sparse decomposition with convolutional Basis Pursuit DeNoising (BPDN) dictionary learning is utilized to get the better compression of the ROI, thus higher perceptual quality and resistance to noise addition attacks. The proposed method solves the drawbacks in the previous methods in the literature. Moreover, the proposed method successes in accurately identifying the distorted region in the gray and color medical image and reconstruct it in very low complexity time.

3. Research Methodology

In this work, SD and SVD are used for assuring the integrity and maintain authenticity. The proposed scheme is decomposed into four main procedures: ROI detection and enhancement procedure, ROI embedding procedure (Algorithm 1), extraction procedure (Algorithm 2), and tamper detection/ correction procedure (Algorithm 3). Fig. 1 shows the block diagram of our proposed system.

A. ROI Detection and Enhancement

The region of interest (ROI) of the medical data is the region that carry the most crucial information. This part is sensitive and should be preserved during transmission via e-healthcare system. Thus, the proposed method automatically separates the ROI and hides it again in the most robust part of the medical image. First, the proper threshold for the binarization of medical image is determined using Otsu's thresholding algorithm [28], followed by automatic thresholding. Then, morphological closing is performed to connect adjacent objects and pruning the edges of the ROI. Next, hole-filling operation is applied to get the final ROI segmentation. Finally, the non-interest region is omitted by assigning the intensities of this region to zeros, thus it does not affect the embedding process and degrade the quality of the watermarked image. Fig. 1 (1) shows the detected ROI in the given medical image.

B. Embedding procedure

The input of this stage is the host medical image and the output is the watermarked image with hidden ROI. Fig. 1 (2) describes the embedding stage of the proposed scheme. First, if the medical image is colored, it is converted into YC_rC_b color space. To make this algorithm more robust, the Y plane is chosen as a host part in the color image as it has higher values than the other components. Then, SVD is applied on the host part (i.e., the grayscale image or Y plane in color image) to get three matrices, i.e., U, S, V. Sparse Decomposition (SD) is also applied on the ROI of the selected part in order to generate the watermark. SD is applied to dramatically decrease the image information and improve payload power using convolutional BPDN learning dictionary [29]. In this form of sparse decomposition, the unstructured dictionary is replaced with a dictionary D with a structure equivalent to convolution with a set of linear filters. The approximation of signal S from its sparsely represented signal S' can be performed by $S \approx D \times S'$, where S can be a whole image rather than just a little image patch. Additive embedding with a predefined scaling factor (i.e., γ=0.5) is then performed to integrate a copy of the medical image ROI in the host part. Then, SVD is utilized again on the resulted watermarked matrix to get three watermarked matrices, i.e., Uw, Sw, Vw. The resulted Uw and Vw matrix are saved as secrete keys (K) used for extraction. The detailed steps and mathematical equations of this procedure are provided in Algorithm 1.

C. Extraction Procedure

To ensure authenticity of the transmitted medical data, the concealed watermark is reconstructed and compared with the received image. To do this, the RGB watermarked image is first transformed into YC_rC_b color space and the

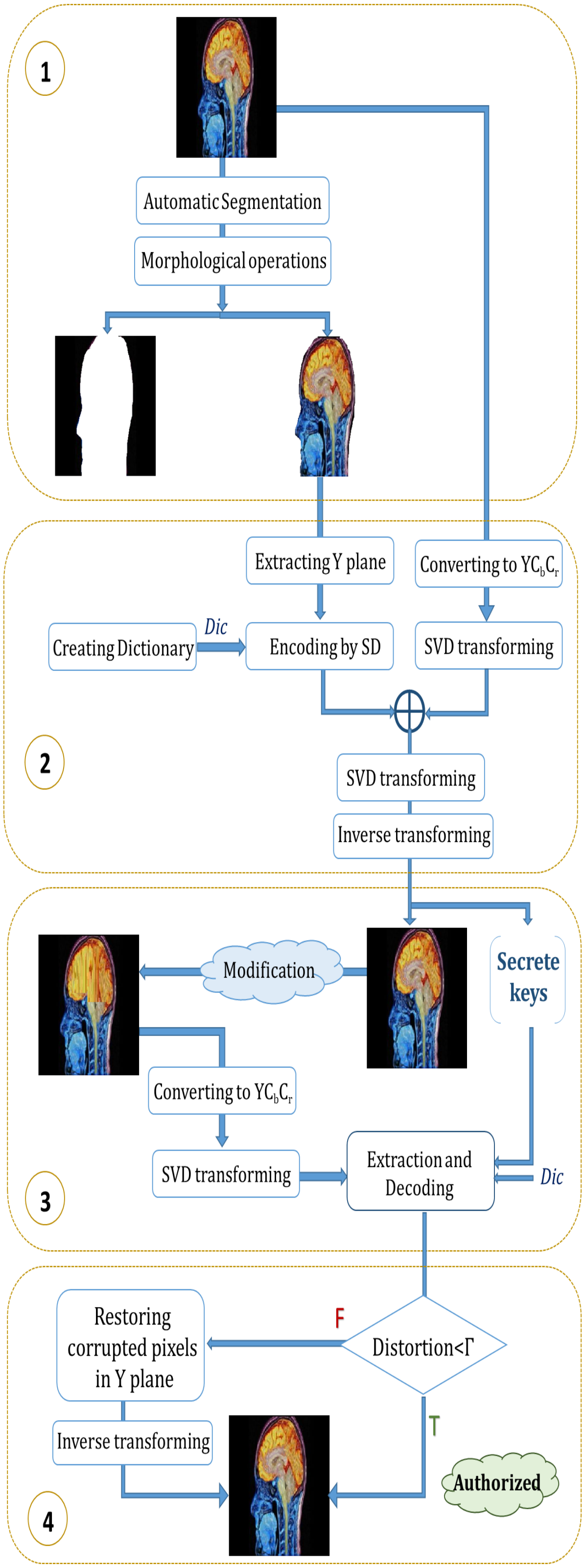


Figure 1. Block diagram of the proposed scheme: (1) localization of ROI, (2) Encoding and embedding procedure, (3) Extraction and decoding procedure, and (4) Detection and correction procedure.

Algorithm 1 The embedding procedure.

```
procedure EMBEDDING
Input: the original color medical image $Im \in R^{M\times N\times O}$
Output: the medical image with hidden ROI $Im' \in R^{M\times N\times O}$ & the security key K
    Begin:
  if color then
      $[YC_bC_r]=YC_bC_r(Im)$
      $Y_o = Y_p = Y$
  end if
  [U Σ V]=SVD ($Y_o$)                                   ▷ SVD
  $\Phi$ = DL(D)                                         ▷ Dictionary learning
  for y do = 1 To N
      X(:, y) = SD ($\Phi$, $Y_p(:,y)$)                   ▷ SD
      $\widehat{\Sigma}$(1:N,y)= $\Sigma(1:N,y)$+ $\gamma$.X(:,y)   ▷ Embedding
  end for
      $[U_x\widehat{\Sigma_x}V_x]$=svd($\widehat{\Sigma}$)
      K= $[U_xV_x]$                                     ▷ Security key generation
      $\widehat{Yo} = [U \times \widehat{\Lambda_h} \times V']$      ▷ Inverse SVD
      I = ($\widehat{Yo}$, $C_b, C_r$)
      Im'= RGB (I)
  end
```

Y channel is separated. Then, SVD is performed on the watermarked part Y. The inverse of embedding process is applied here to get the hidden part, as shown in Fig. 1 (3). Specifically, the singular matrix of the watermarked Y plane is combined with the secrete key (K) from embedding stage. The sparse matrix is reconstructed by subtraction decoding function using the same scaling factor γ used for embedding. The hidden ROI is then reconstructed by multiplying the extracted sparse matrix with the learning dictionary D. The inputs, outputs, and details of this procedure are provided in Algorithm 2.

Algorithm 2 The extraction procedure.

```
procedure EXTRACTION
Input: the received medical image with hidden message $Im'_R \in \mathbb{R}^{M\times N\times O}$ & dictionary $\Phi$ & the security key K
Output: the extracted Y pattern $P_{ex} \in \mathbb{R}^{M\times N}$
    Begin:
  if color then
      $[Y_RC_bC_r]=YC_bC_r(Im'_R)$
  end if
  $[U_R\Sigma_RV_R]$=SVD ($Y_R$)
  $Y_Ex = K(M,N_1) \times \Sigma_R \times K(M,N_2)^T$      ▷ Reconstruct $Y_{Ex}$
  $X_R = (Y_{Ex} - \Sigma_R).\gamma$                        ▷ Extraction
  $P_{ex} = \Phi \times X_R$                               ▷ Reconstruct the hidden ROI
end
```

D. Tamper Detection and Correction Procedure

The aim of this phase is to identify, localize, and restore the distorted region of the received image based on the extracted concealed watermark from the previous procedure. The distorted region can be identified by calculating the absolute difference between the reconstructed part P_ex and the corresponding part of the modified image P_d. If the computed difference at particular (x, y) coordinate is higher than a specific threshold value τ, then, this medical image is considered as unauthentic image and restored by replacing the modified pixels by the corresponding values in the extracted pattern, as illustrated by the following equation:

$$P_R(x,y) = \begin{cases} P_{ex}(x,y), |P_{ex}(x,y) - P_d(x,y)| > \tau \\ P_d(x,y), Otherwise \end{cases} \tag{3}$$

Afterwards, the restored luminance part P_R is combined with the other image Chroma planes. The resulted image is finally transformed back into the RGB color space. Further descriptions and details of this procedures are illustrated by Algorithm 3.

Algorithm 3 The tamper detection and correction procedure

```
procedure EMBEDDING
Input: the received image $Im'_R \in \mathbb{R}^{M\times N\times O}$ & the extracted pattern $P_{ex} \in \mathbb{R}^{M\times N}$
Output: the corrected medical image $Im_c \in \mathbb{R}^{M\times N\times O}$
    Begin:
  if color then
      $[Y_R\ C_bC_r]=YC_bC_r\ (Im'_R)$
  end
  for I do = 1: M
      for j do = 1: N
          $d(x,y) = |P_{ex}(x,y) - Y_R(x,y)|$
          if $d(x,y) > \tau$ then
              mask(x,y)=1
              $Y_R$(x,y) = $P_{ex}$(x,y)      ▷ Correcting the distorted pixels
          end if
      end for
  end for
  If mask=0 → Authentic Medical Image
  $I_c$ = ($Y_R$, $C_b$, $C_r$)
  $Im_c$=RGB ($I_c$)
end
```

4. Results and Discussion

The experimental analysis is performed using different medical images of 256 × 256 dimensions from ”Mid-Pix” which is an open-access database for medical images (https://medpix.nlm.nih.gov). Also, some images from a medical search engine, i.e., Openmd (https://openmd.com/), are used. These medical images have different details for different human body portions. The simulation tests were implemented using MATLAB 2020a operated on a Windows 10 machine with a 2.40 GHz Intel Core i5 processor and 8 GB RAM. Sparse decomposition is performed using the “SParse Optimization Research Code” (SPORCO) [30], which is an open-source library for solving optimization problems with sparsity regularization, dictionary learning, for both standard and convolutional forms of sparse coding.

TABLE I. Summary of different measures used for performance evaluation of the proposed method.

Eq#	Metric name	Mathematical expression	Definition	Optimal value
1	Normalized Correlation (NC)	$NC = \frac{\sum_x \sum_y W \times W'}{\sqrt{\sum_x \sum_y W^2 \times \sum_x \sum_y W'^2}}$	*W*: the original *W'*: extracted watermark w & h: image width and high	1: optimal case -1: no similarity
2	Normalized Absolute Error (NAE)	$NAE = \frac{\sum_{x=1}^{w} \sum_{y=1}^{h} \lvert W(x,y) - W'(x,y) \rvert}{\sum_{x=1}^{w} \sum_{y=1}^{h} \lvert W(x,y) \rvert}$	*W* & *W'* are the original and extracted watermark	0: optimal case 1: no similarity
3	Mean Square Error (MSE)	$MSE = \frac{1}{\mathrm{M} \times \mathrm{N}} \sum_{x=1}^{M} \sum_{y=1}^{N} (f\,(x,y) - g\,(x,y))^2$	f and g are two images of size M × N. $f\,(x,y)$ and $g\,(x,y)$ are the intensities of the *(x,y)* coordinate.	0: optimal case 1: no similarity
4	"Peak Signal-to-Noise Ratio" (*PSNR*)	$PSNR = 10 \times log_{10} \frac{255^2}{MSE}$	f and g are two M × N images	>40 (dB) indicates high similarity
5	"Structural Similarity Index Measure" (SSIM)	$SSIM = (l(f_B, g_B)^{\alpha}.c(f_B, g_B)^{\beta}.s(f_B, g_B)^{\gamma}) \times 100\%$	f_B and g_B are any two image blocks. *SSIM* compares a luminance (***l***), contrast (***c***), and structure (***s***) planes. $\alpha, \beta, \gamma \geq 1$ are the importance weight.	>0.95 indicates high similarity
6	Weighted peak signal-to-noise ratio (wPSNR)	$wPSNR = 10 \times log_{10} \frac{255^2}{\lvert\lvert NVF^2 \rvert\rvert}$	NVF: Noise Visibility function calculated by $NVF = \frac{1}{1+\sigma(i,j)^2}$ $\sigma(i,j)^{\,2}$ represents the local variance.	>40 (dB) indicates high similarity
7	Unified Average Changing Intensity (UACI)	$UACI = \frac{1}{\mathrm{M} \times \mathrm{N}} \sum_{x=1}^{M} \sum_{y=1}^{N} \frac{\lvert f\,(i,j) - g\,(i,j) \rvert}{255 - 0} \times 100\%$	f and g are two M × N images	>33.46%
8	Number of Pixels Change Rate (NPCR)	$NPCR = \frac{1}{\mathrm{M} \times \mathrm{N}} \sum_{x=1}^{M} \sum_{y=1}^{N} \lvert \mathrm{sign}(f\,(x,y) - g\,(x,y)) \rvert \times 100\%$	Sign(.) is the sign function determined based on the difference between the two images *f*-*g*=Δ, as follows: $sign = \begin{cases} 1\ if\ \Delta > 0 \\ 0\ if\ \ \Delta = 0 \\ -1\ if\ \Delta < 0 \end{cases}$	>99.609%

A. Performance Evaluation Functions

In this paper, several evaluation functions are used. Table I summarizes the measures used and description of each one. Five measures are used to evaluate our embedding phase: Peak Signal-to-Noise Ratio ($PSNR$), Structural Similarity Index Measure ($SSIM$), Weighted Peak Signal-to-Noise Ratio ($wPSNR$), Unified Average Changing Intensity ($UACI$), and Number of Pixels Change Rate ($NPCR$), which are computed as shown in Eq. (4-8) in Table I, respectively. To calculat the $PSNR$, Mean Square Error (MSE) is first calculated based on Eq. (3) in Table I. A $PSNR$ value of more than 40dB indicates that the watermarked image falls within normal degradation limits and the hidden pattern is almost invisible [31], whereas SSIM above 0.95 are generally considered to be visually unnoticeable to the human eye [32], [33].

The traditional $PSNR$ assesses distortion on the host data without taking the human perception into account. Thus, the weighted Peak signal-to-Noise Ratio ($wPSNR$) is also computed for our imperceptibility test. $PSNR$ and $wPSNR$ are generally similar for the smooth areas, but for the high texture areas, $PSNR$ is lower than $wPSNR$ as the noise visibility function is almost zero [34].

The $UACI$ and $NPCR$ measures are computed in order to determine the proportion of distinct intensities between two images. $UACI$ defines the average intensity of dissimilarities between two images, whereas $NPCR$ shows the pro-

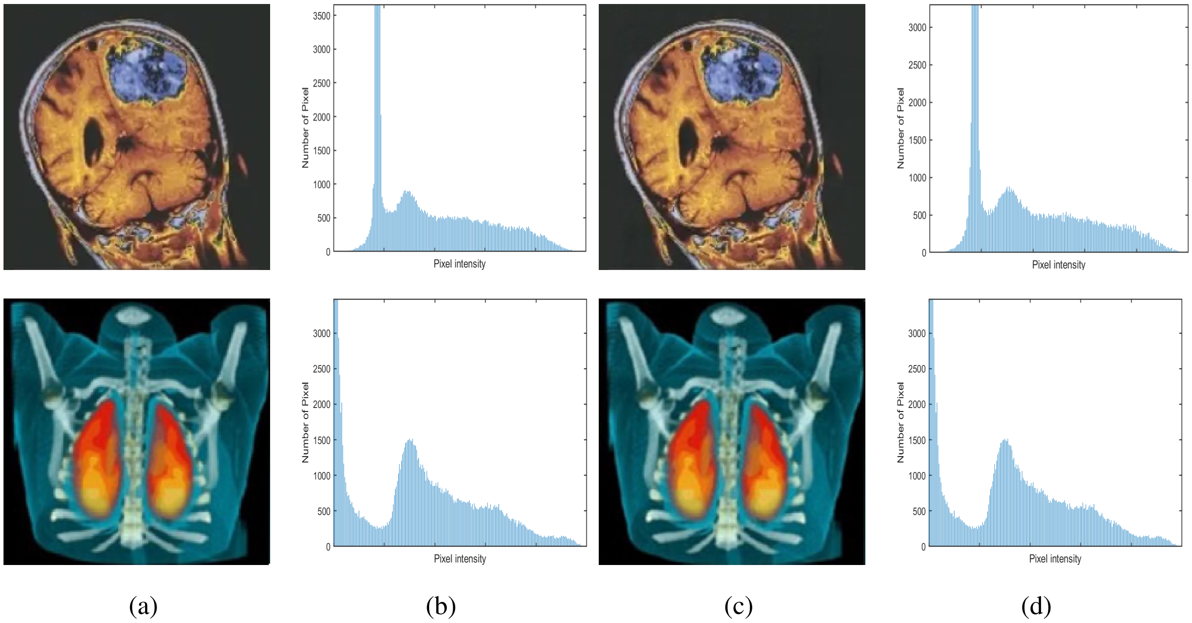


Figure 2. Column (a) displays two color medical images with the corresponding histograms displayed in column (b). Column (c) represents the carrier images with the corresponding histograms of the selected layer displayed in column (d).

portion of distinct intensities between two digital images. Obtaining greater $UACI$ and $NPCR$ values is recommended for improving security (resistance to differential attacks). The optimal percentage value for UACI is 33.4635% while the optimal percentage value for NPCR is 99.6094% [35], [36].

For robustness evaluation, Normalized Correlation (NC) and Normalized Absolute Error (NAE) are computed to assess the similarities between the original and reconstructed part. Equations (1) and (2) in Table I give the mathematical expressions of theNC and NAE measures. The satisfying watermarking security, i.e., resistance to attacks, is associated with high NC values (nearly 1), and low NAE values (nearly 0).

B. Evaluation of Embedding Procedure

This section assesses the quality of the medical images after self-embedding. Table II displays the $PSNR$, $SSIM$, $wPSNR$, $UCAI$, and $NPCR$ for six sample of medical images used in our experiments. All $PSNR$ values are above 48 (dB) with an average of around 50, and all $wPSNR$ values are greater than 52dB. Also, all $SSIM$ values are near 1, even though the high embedding capacity, i.e., the same size of the host medical image. In addition, all $UACI$ values are near to the optimal value of around 33%. Our scheme has also a high $NPCR$ that are closer to the theoretical percentage values.

These outcomes demonstrate a high level of imperceptibility achieved by our embedding procedure owing to utilizing the dimensionality reduction tool, i.e., sparse decomposition, so just a small number of non-zero values

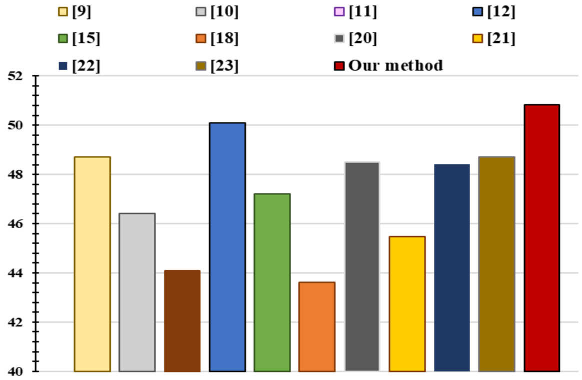


Figure 3. PSNR comparison between our SD-based method and the existing methods.

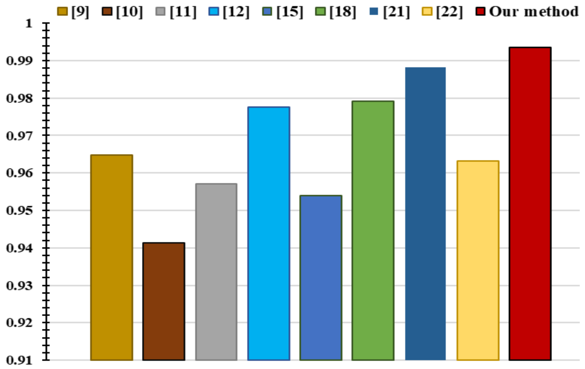


Figure 4. SSIM comparison between our SD-based method and the existing methods.

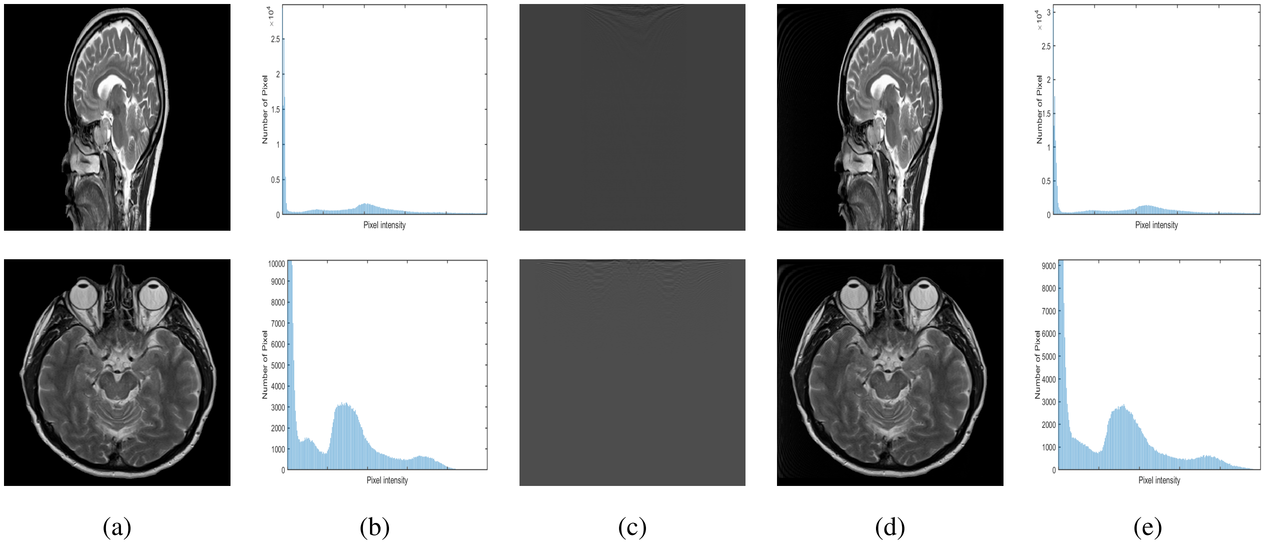


Figure 5. Column (a) displays two medical images used in our experiments with the corresponding histogram in column (b). Column (c) represents the ciphered region, whereas columns (d) and (e) represent the constructed hidden region with the corresponding histogram, respectively.

TABLE II. The performance of embedding procedure.

Images.	PSNR	SSIM	wPSNR	UACI	NPCR
Med-Img1	48.47	0.9936	55.44	32.38	99.99
Med-Img2	54.36	0.9997	54.38	32.94	98.76
Med-Img3	50.03	0.9902	57.22	29.04	99.55
Med-Img4	50.07	0.9944	52.62	32.47	94.40
Med-Img5	48.71	0.9923	60.81	29.76	99.99
Med-Img6	50.25	0.9934	52.99	32.24	95.71

modify the weight of the host medical data. Moreover, the visual inspection of the invisibility aspect of our method is confirmed by Fig. 2. The figure displays samples of color medical images used for our experiments (a), the corresponding watermarked images with the hidden ROI in (c), and their histograms shown in (b) and (d). As shown in the figure, our watermarked images have a high quality, and it is quite hard for the human eye to identify the concealed information. In addition, there is high invisibility proven by the similarities between the histograms of the host and carrier images.

For further evaluation, the performance of our system is evaluated in comparison with the performance of other current methods in term of $PSNR$ and $SSIM$, and the results are displayed in Fig. 3 and Fig. refFig:SSIM, respectively. As the figures demonstrate, the average $PSNR$ and $SSIM$ values for our proposed method are much higher than other methods because of utilizing learned sparse decomposition where only a few numbers of elements enough to reconstruct the hidden ROI instead of all elements. Finally, the obtained results show that it quite challenging for the human to detect and identify the concealed data, where the image modifications look as random noise frequently present in the digital images.

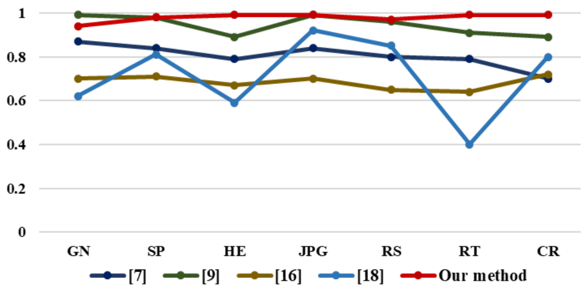


Figure 6. Comparison of the robustness of our proposed scheme and existing schemes under some attacks.

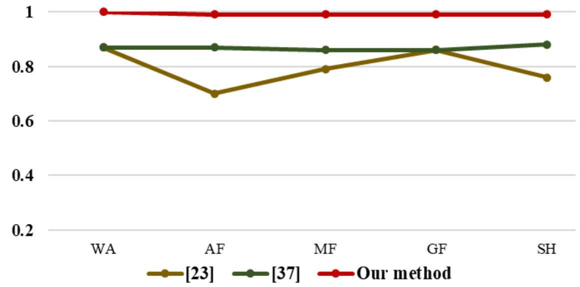
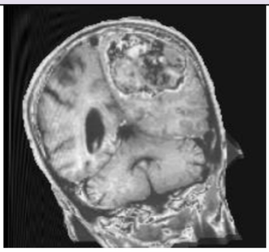
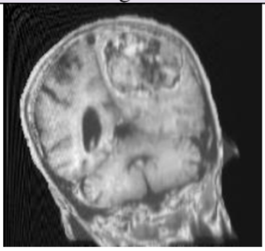
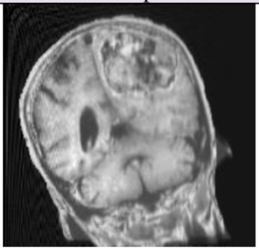
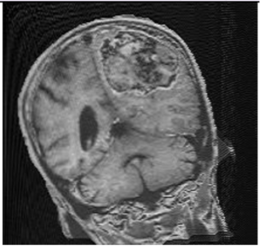
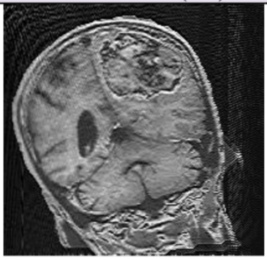
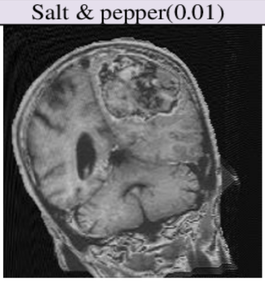
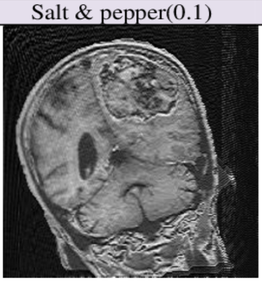
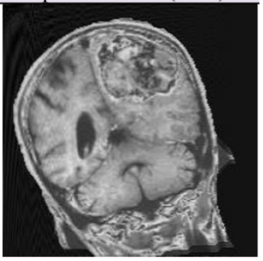
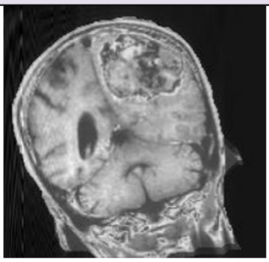
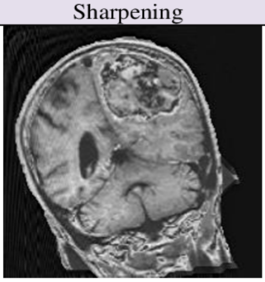
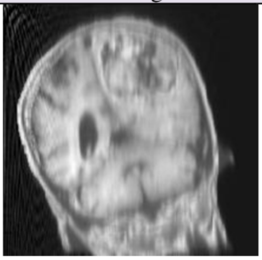
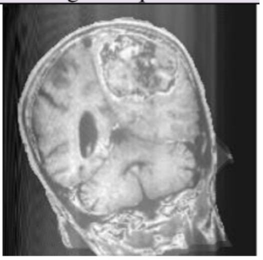
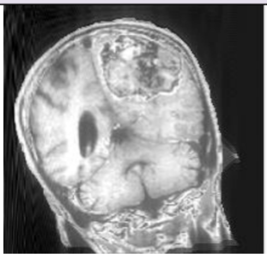
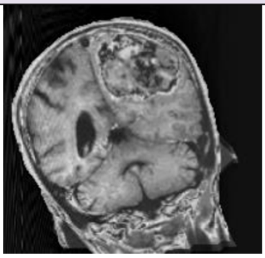
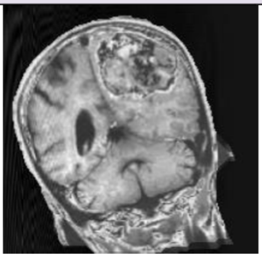
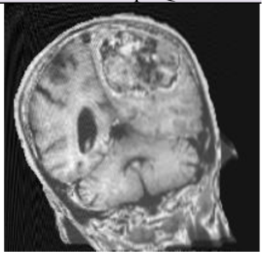
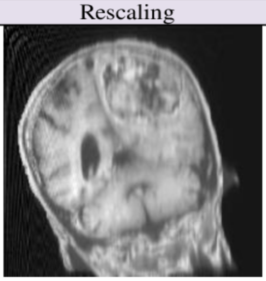
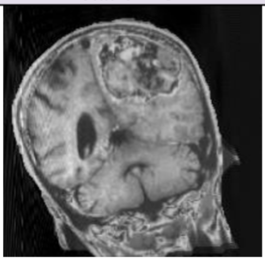
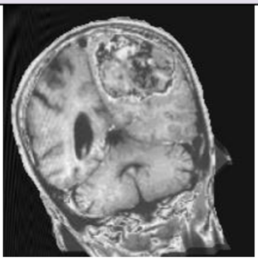
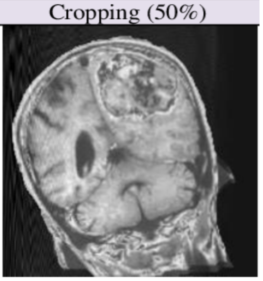


Figure 7. Comparison of the robustness performance against filtering attacks.

C. Evaluation of Extraction Procedure

For extraction procedure evaluation, Fig. 5 shows two grayscale medical images used as watermarks and the corresponding histogram in (a) and (b), respectively. The encrypted version of the watermark with SD is displayed in Fig. 5 (c) and the extracted watermark with the corresponding histogram are presented in (d) and (e), respectively.

TABLE III. Samples of extracted tampered pattern under various attacks.

Attacks	No attack	Average filter	Gaussian Lowpass filter	Gaussian noise (0.01)
Extracted hidden part				
NC/ NAE	**0.9938/ 0.0233**	**0.9819/ 0.0789**	**0.9838/ 0.0733**	**0.9569/ 0.1425**
Attacks	Gaussian noise (0.1)	Salt & pepper(0.01)	Salt & pepper(0.1)	Speckle Noise(0.01)
Extracted hidden part				
NC/ NAE	**0.8817/ 0.2654**	**0.9799/ 0.0913**	**0.9082/ 0.2233**	**0.9912/ 0.0483**
Attacks	Gamma correction-0.1	Sharpening	Blurring	Histogram equalization
Extracted hidden part				
NC/ NAE	**0.9750/ 0.9042**	**0.9854/ 0.1123**	**0.9632/ 0.1225**	**0.9408/ 0.6616**
Attacks	Gamma correction-0.5	Soft threshold	Hard threshold	JPEG comp. QF=30
Extracted hidden part				
NC/ NAE	**0.9850/ 0.3432**	**0.9879/ 0.3418**	**0.9926/ 0.0505**	**0.9891/ 0.0548**
Attacks	Rescaling	Rotation 25 °	Rotation 90 °	Cropping (50%)
Extracted hidden part				
NC/ NAE	**0.9730/ 0.1019**	**0.9860/ 0.0828**	**0.9938/ 0.0232**	**0.9867/ 0.1065**

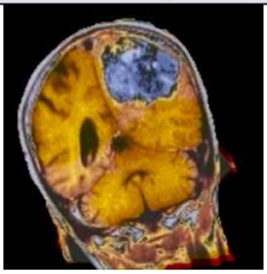
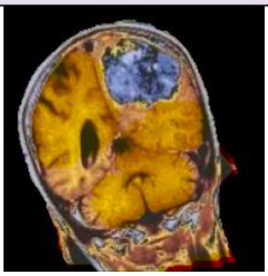
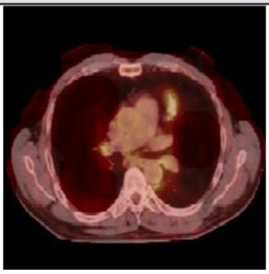
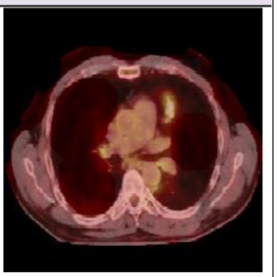
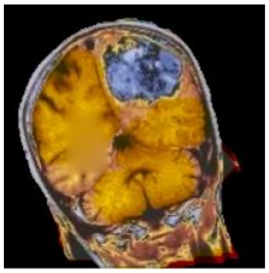
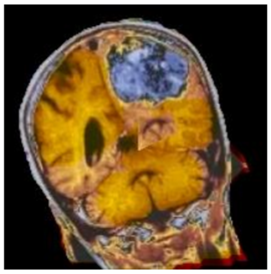
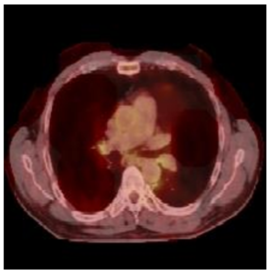
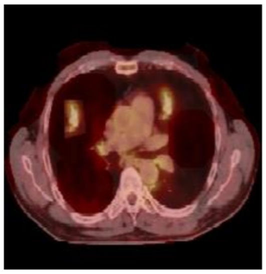
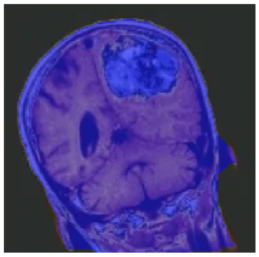
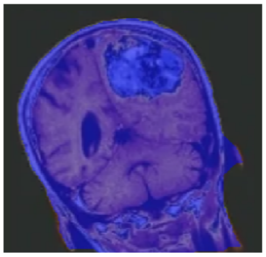
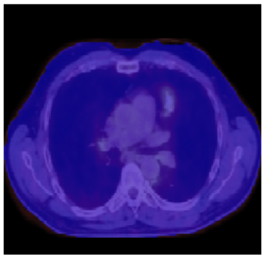
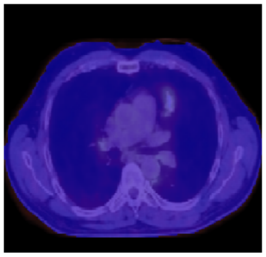
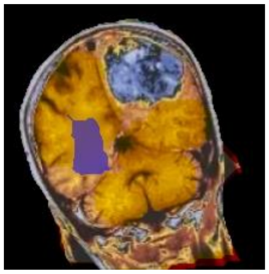
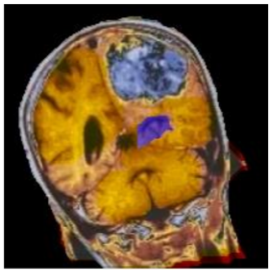
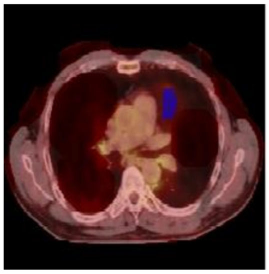
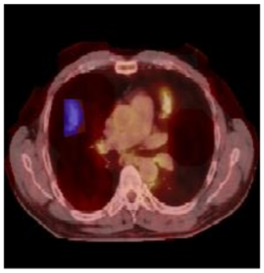
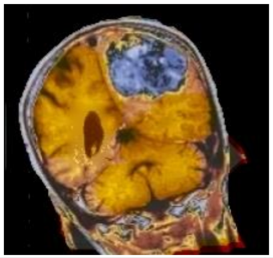
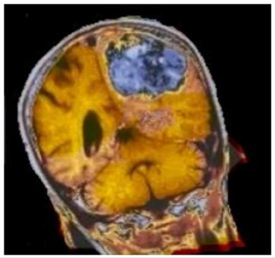
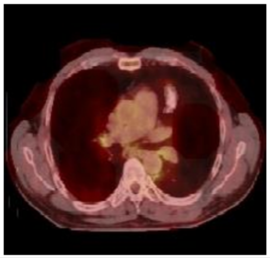
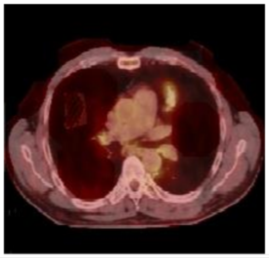
Attacks	Removing	Addition	Removing	Addition
Original Medical image				
Distorted image				
Detected ROI				
Detected tamper shown by blue region				
Restored image				
PSNR	43.0039	41.8147	46.3954	45.4324
SSIM	0.9271	0.9237	0.9618	0.9560
Time	0.69s	0.61 s	0.57 s	0.46

Figure 8. The detection and correction results of the proposed method under tamper addition and region removing attacks, where row 1 and 2 display the original and distorted images, where row 3 and row4 show the detected ROI and tampered area in the ROI highlighted by blue region, respectively. The corresponding corrected images by our method are shown in row 4. The *PSNR*, *SSIM*, and time measures for our correction procedure are presented in the rest rows, respectively.

As shown, the encrypted part has no similarity with the origin part, however, it is successfully used to reconstruct the hidden watermark with high perceptual quality. Another advantage of our method is that it offers a natural encryption technique to protect the hidden watermark (see Fig. 5 (c)).

In watermarking applications, the reliable method has to be capable of retrieving the hidden data even when the image is seriously destroyed. To evaluate the reliability, the proposed scheme has been evaluated under some serious attacks, including sharpening, noise adding, cropping, gamma correction, compression. Fig. 6 represents *NC* comparison between our method and some existing techniques in terms of noise addition, translation, and compression attacks, i.e., Gaussian noise with variance 0.01 (GN), impulse noise (SP), histogram equalization (HE), JPEG compression with QF=30 (JPG), rescaling 0.5 (RS), rotation with angle 45 ° (RT), and 25% cropping (CR). Based on the results in the figure, our proposed method significantly surpasses other methods, and it is scale- and rotation-invariant.

Moreover, the proposed scheme is robust against serious geometric and noise attacks. The proposed method also shows a high resistance to filtering attacks. Fig. 7 shows a comparison between our method and [23], [37] under different filtering attacks, including with [3 × 3] Average Filter (AV), Gaussian Low-pass (GF), Median Filter (MF), and Sharpening (SH). The robustness performance of our proposed scheme is much higher than other schemes for all kind of filtering attacks.

For further evaluation, Table III illustrates the qualitative and quantitative results. According to the table, the *NC* value of proposed system for the watermarked medical image without attack is always near to 1. Also, the original and extracted watermarks have a very high degree of similarity, with an average *NC* of 97% and an average NAE of 0.19258. Even with high density noise addition, the hidden watermark is still extractable with good quality, i.e., *NC* is greater than 0.88. These outcomes prove that the proposed technique offers a good robustness even though the distortion in the received medical image is high.

D. Evaluation of tamper detection and correction

The aim of this test is to confirm the ability of our method in providing a secure transmission of the medical data against intentional attacks. The proposed scheme embeds crucial information of medical data, represented by the ROI, in the most stable part of the color image. On the receiver side, the hidden part is restored again and compared with the corresponding part of the received image to confirm the integrity and authenticity. The authenticity of its origin is verified since the origin copy of the crucial region is embedded with the received image.

Fig. 8 evidences the efficiency of the proposed method in detection, localizing and correction intentional attacks, such as adding tamper or removing a part of the medical images. The $PSNR$ and $SSIM$ are also computed to measure the similarity between original and reconstructed medical image for further evaluation. The proposed method shows a high accuracy in localizing the tampered area, and high integrity represented by the ability to restore corrupted images in low complexity time as shown in the table.

Overall, experimental result analysis proves the high imperceptibility of the updated medical image, where the hidden patter does not affect the image quality. Moreover, it is demonstrated that our scheme is robust against various attacks, including geometric attacks, JPEG compression, and signal processing attacks, such as filtering, noise adding, blurring, contrast enhancement. In addition, the analysis carried out to prove the effectiveness of the proposed system in ensuring authenticity proves that the proposed method successes in achieving a high security level and determining authentic and unauthentic image.

Finally, the provided results have demonstrated that our proposed scheme successes in restoring the tamper area, achieving both invisabilty and robustness aspects, and offering a good encryption as well as a high level of compression. The proposed method has the capability to localize and reconstruct the tampered area with high similarity measures between the host and reconstructed image. It is also proven that our scheme outperforms the current methods with regards to perceptual quality, robustness, detection and correction performance.

5. Conclusions

This paper introduces a new hiding scheme to verify the integrity of grayscale and color medical images based on two decomposition and dimensionality reduction techniques, singular value decomposition and learning sparse decomposition, i.e., singular value decomposition and learning sparse decomposition. Unlike most of existing methods in the literature, the ROI in the medical image is identified by automatic segmentation and morphological operations. Also, dictionary learning is used which is particularly helpful in the context of input image adaptation. The proposed system is assessed using different kinds of gray and color medical images under different attacks. A performance comparison is also made between our scheme and some current schemes.

The proposed system presents a high performance with regards to invisibility, robustness, and security, which are the most necessary requirement of reliable watermarking systems. It has been also demonstrated that our proposed technique can be successfully used to guarantee the security and confidentiality of both gray and color medical images. It is found that involving SD with convolutional Basis Pursuit DeNoising (BPDN) dictionary learning in our method offers many utilities such as invisibility, resistance the noise, a high compression and a very strong encryption. As a future work, genetic algorithm can be utilized to dynamically choose the most suitable scaling factor used in the embedding and extraction procedure. Also, the recovery capacity can be further improved.

References

[1] C.-C. Chang, P. Tsai, and C.-C. Lin, “Svd-based digital image watermarking scheme,” *Pattern Recognition Letters*, vol. 26, no. 10, pp. 1577–1586, 2005.

[2] H. Lee, A. Battle, R. Raina, and A. Ng, “Efficient sparse coding algorithms,” *Advances in neural information processing systems*, vol. 19, 2006.

[3] A. Tareef and A. Al-Ani, “A highly secure oblivious sparse coding-based watermarking system for ownership verification,” *Expert Systems with Applications*, vol. 42, no. 4, pp. 2224–2233, 2015.

[4] A. K. Jabbar, A. T. Hashim, and Q. F. Hassan, “Medical image authentication by combining hash signature and watermarking based on frequency domains,” in *Journal of Physics: Conference Series*, vol. 1963, no. 1. IOP Publishing, 2021, p. 012039.

[5] W. El-Shafai, I. Almomani, A. Ara, and A. Alkhayer, “An optical-based encryption and authentication algorithm for color and grayscale medical images,” *Multimedia Tools and Applications*, vol. 82, no. 15, pp. 23 735–23 770, 2023.

[6] F. Alqahtani, M. Amoon, and W. El-Shafai, “A fractional fourier based medical image authentication approach.” *Computers, Materials & Continua*, vol. 70, no. 2, 2022.

[7] S. A. Parah, J. A. Sheikh, F. Ahad, N. A. Loan, and G. M. Bhat, “Information hiding in medical images: a robust medical image watermarking system for e-healthcare,” *Multimedia Tools and Applications*, vol. 76, pp. 10 599–10 633, 2017.

[8] P. Sivananthamaitrey and P. R. Kumar, “Optimal dual watermarking of color images with swt and svd through genetic algorithm,” *Circuits, Systems, and Signal Processing*, vol. 41, pp. 224–248, 2022.

[9] K. Swaraja, K. Meenakshi, and P. Kora, “An optimized blind dual medical image watermarking framework for tamper localization and content authentication in secured telemedicine,” *Biomedical Signal Processing and Control*, vol. 55, p. 101665, 2020.

[10] E. Gul and S. Ozturk, “A novel pixel-wise authentication-based self-embedding fragile watermarking method,” *Multimedia Systems*, vol. 27, no. 3, pp. 531–545, 2021.

[11] S. Prasad and A. K. Pal, “A tamper detection suitable fragile watermarking scheme based on novel payload embedding strategy,” *Multimedia Tools and Applications*, vol. 79, no. 3-4, pp. 1673–1705, 2020.

[12] S. Gull, N. A. Loan, S. A. Parah, J. A. Sheikh, and G. M. Bhat, “An efficient watermarking technique for tamper detection and localization of medical images,” *Journal of ambient intelligence and humanized computing*, vol. 11, pp. 1799–1808, 2020.

[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, pp. 12 881–12 903, 2021.

[14] A. Soualmi, A. Alti, and L. Laouamer, “An imperceptible watermarking scheme for medical image tamper detection,” *International Journal of Information Security and Privacy (IJISP)*, vol. 16, no. 1, pp. 1–18, 2022.

[15] H. Rhayma, A. Makhloufi, H. Hamam, and A. B. Hamida, “Semi-fragile watermarking scheme based on perceptual hash function (phf) for image tampering detection,” *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26 813–26 832, 2021.

[16] T. Naheed, I. Usman, T. M. Khan, A. H. Dar, and M. F. Shafique, “Intelligent reversible watermarking technique in medical images using ga and pso,” *Optik*, vol. 125, no. 11, pp. 2515–2525, 2014.

[17] T. Sun, X. Wang, D. Lin, R. Bao, D. Jiang, B. Ding, and D. Li, “Medical image security authentication method based on wavelet reconstruction and fractal dimension,” *International Journal of Distributed Sensor Networks*, vol. 17, no. 4, p. 15501477211014132, 2021.

[18] A. K. Singh, “Robust and distortion control dual watermarking in lwt domain using dct and error correction code for color medical image,” *Multimedia Tools and Applications*, vol. 78, pp. 30 523–30 533, 2019.

[19] R. Eswaraiah and E. S. Reddy, “Medical image watermarking technique for accurate tamper detection in roi and exact recovery of roi,” *International journal of telemedicine and applications*, vol. 2014, pp. 13–13, 2014.

[20] H. L. Khor, S.-C. Liew, and J. M. Zain, “Region of interest-based tamper detection and lossless recovery watermarking scheme (roi-dr) on ultrasound medical images,” *Journal of digital imaging*, vol. 30, pp. 328–349, 2017.

[21] A. R. Gottimukkala, A. Pradhan, S. D. Kosuru, and G. Swain, “Image tamper detection and correction based on mean pixel value and logistic map,” in *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1. IEEE, 2023, pp. 1071–1076.

[22] N. E. H. Golea, K. E. Melkemi, and A. Behloul, “Zero-bit fragile watermarking for medical image tamper detection and recovery using rs code and lifting wavelet transform,” *The Imaging Science Journal*, vol. 69, no. 5-8, pp. 334–349, 2021.

[23] H. S. Alshanbari, “Medical image watermarking for ownership & tamper detection,” *Multimedia tools and applications*, vol. 80, pp. 16 549–16 564, 2021.

[24] F. Ernawan, A. Aminuddin, D. Nincarean, M. F. Ab Razak, and F. Ahmad, “Three layer authentications with a spiral block mapping to prove authenticity in medical images,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, 2022.

[25] Z. Zhang, W. Xiao, T. Liu, Y. Li, S. Jin, F. Li, and H. Wang, “A reversible image watermarking algorithm for tamper detection based on sift,” *Multimedia Tools and Applications*, pp. 1–22, 2023.

[26] R. E. Arevalo-Ancona and M. Cedillo-Hernandez, “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, 2023.

[27] S. Bhalerao, I. A. Ansari, and A. Kumar, “A reversible medical image watermarking for roi tamper detection and recovery,” *Circuits, Systems, and Signal Processing*, vol. 42, no. 11, pp. 6701–6725, 2023.

[28] N. Otsu *et al.*, “A threshold selection method from gray-level histograms,” *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[29] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decompo-

sition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[30] B. Wohlberg, "Sporco: A python package for standard and convolutional sparse representations," in *Proceedings of the 15th Python in Science Conference, Austin, TX, USA*, 2017, pp. 1–8.

[31] T. Barhoom and W. Saqer, "Steganography algorithm within 2-lsbs with indicatorsbased randomness," *International Journal of Computing and Digital Systems*, vol. 6, no. 05, pp. 271–276, 2017.

[32] S. A. Kasmani and A. Naghsh-Nilchi, "A new robust digital image watermarking technique based on joint dwt-dct transformation," in *2008 third international conference on convergence and hybrid information technology*, vol. 2. IEEE, 2008, pp. 539–544.

[33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[34] R. Eswaraiah and E. Sreenivasa Reddy, "Robust medical image watermarking technique for accurate detection of tampers inside region of interest and recovering original region of interest," *IET image Processing*, vol. 9, no. 8, pp. 615–625, 2015.

[35] G. Chen, Y. Mao, and C. K. Chui, "A symmetric image encryption scheme based on 3d chaotic cat maps," *Chaos, Solitons & Fractals*, vol. 21, no. 3, pp. 749–761, 2004.

[36] P. T. Akkasaligar and S. Biradar, "Selective medical image encryption using dna cryptography," *Information Security Journal: A Global Perspective*, vol. 29, no. 2, pp. 91–101, 2020.

[37] A. A. Mohammad, A. Alhaj, and S. Shaltaf, "An improved svd-based watermarking scheme for protecting rightful ownership," *Signal Processing*, vol. 88, no. 9, pp. 2158–2180, 2008.

Afaf Tareef received a B.Sc. degree in computer science from Mutah University, Jordan in 2008, an M.Phil. degree from the University of Jordan in 2010, and a Ph.D. degree from the University of Sydney, Australia in 2017. She is currently an Assistant Professor in the Faculty of Information Technology at Mutah University, Jordan. She has many publications in several international conferences and journals. Her research interests include image processing and medical image analysis.