A New Medical Image Watermarking Technique with Finer Tamper Localization

Arda Ustubioglu¹ · Guzin Ulutas¹

Published online: 27 February 2017
© Society for Imaging Informatics in Medicine 2017

Abstract Medical imaging and information management systems require transmission of medical images over the Internet. Many image watermarking techniques have been proposed in recent years to ensure the integrity and authenticity of medical images transferred over insecure networks. In this work, we propose a new medical image watermarking technique to detect tampered regions on medical images with finer accuracy by authenticating 4×4 blocks and without restricting region of interest (ROI) size. The proposed method can mark a 4×4 pixel block if it has even one tampered pixel, while similar methods (which have no ROI size restriction) mark 8×8 , 16×16 , and 40×40 pixel blocks. Modified difference expansion (MDE) and least significant bit (LSB) embedding techniques are used together first in the literature by the method to embed authentication bits into corresponding blocks. The method uses a small 4×4 window to mark the tampered region. Experimental results show that the proposed method detects tampered regions on medical images with high accuracy and can be used by all medical image modalities. The results also indicate that the method has finer accuracy and no ROI size restriction compared to similar works reported in the literature.

Keywords Medical image watermarking · Tamper localization · Difference expansion

Introduction

Recent developments in information and communication technologies have changed medical imaging and information management systems at the hospitals considerably. Digital storage media replaced hard copy medical image on films, and telemedicine systems have enabled medical image and electronic patient record (EPR) exchange among physicians worldwide. Medical information systems (MISs) provide sophisticated healthcare delivery systems by using the digital imaging and communications in medicine (DICOM) standard. HIS provides remote access to medical records, transmits them over the Internet, and eases diagnose by telediagnosis.

However, medical image and data transmission over the Internet is not secure unless strong encryption techniques are used. Best effort routing strategy of many routers makes the data vulnerable to interception and may result in tampering of the medical images during transmission over the Internet. Modification on medical images could result in misdiagnosis. Three requirements must be satisfied by any medical information systems to make sure that medical images are intact: confidentiality, authenticity, and integrity [1]. Confidentiality imposes that only entitled users can access to the transmitted medical image. Authenticity shows that the medical image belongs to the correct patient and comes from the known source. The last requirement proves that the medical image is intact.

Two techniques are proposed in the literature to deal with the three requirements: cryptography and watermarking. Symmetric and asymmetric encryption algorithms can be used in medical information systems. But, the system cannot verify the integrity and authenticity of the medical image if the digital signature is lost or modified. The latter technique embeds information called a watermark into the image and it ensures the three requirements given above. Watermarking has been

✉ Arda Ustubioglu
ardaustubioglu@ktu.edu.tr

¹ Computer Engineering Department, Karadeniz Technical University, Trabzon, Turkey

used by the medical information systems and many methods have been proposed recently.

It is possible to classify medical image watermarking techniques depending on their characteristics. One of the classifications is their watermark embedding techniques. In this context, we grouped the methods in the literature into three sections: irreversible methods, reversible methods, and region-based methods. The methods in the first category modify the medical image with an irreversible manner during watermark embedding. Distortions on the medical images cannot be acceptable because it can cause misdiagnosis. The methods in the second category reconstruct the original medical image after watermark extraction and hence do not affect diagnosis. The methods in the last category split medical images into two regions: region of interest (ROI) and region of non-interest (RONI). These methods use different watermarking techniques for two regions. While reconstruction of ROI after watermark extraction procedure is important, any irreversible watermarking technique can be used for RONI. The most important problem with these methods is to know the exact location of ROI. Either clinician determines the exact location of ROI or side information that contains the coordinates of ROI region must be embedded as a payload. Watermarking technique can also extract ROI region automatically. Region-based methods can be impractical when the technique is used as a general solution at the HIS.

An outline of the medical image watermarking field that uses various techniques to embed watermark data and utilize various functions to detect tampered regions is given below in chronological order.

A method proposed in 2005 by Woo et al. uses more than one watermarking technique to watermark medical images [3]. Pixels in the border of the image are modified to carry patient record information. Robust watermarking technique is used in this region to hide EPR. LSBs of the other pixels except from border pixels are modified to accommodate visible watermark information. Fragile watermarking technique used in ROI makes possible to detect tampered regions inside this region. However, this method is not reversible and the medical image is modified irrecoverably after embedding procedure.

The authors in [4] proposed two authentication methods by using discrete cosine transform (DCT) and modulo operator. In the first scheme, hash value of the image is used for authentication purposes and jpeg compressed version of the image is utilized for recovery. However, compression rate must be high in order not to reduce embedding capacity. High compression rate results in poor reconstructed image quality. The second scheme only considers the ROI region and generates authentication/recovery information from ROI. Lower compression rate can be chosen and therefore ROI can be recovered with better visual quality. The method described in [4] is not reversible because the pre-processing stage modifies the

original image irrecoverably. Also, the method authenticates whole ROI with a single hash value which reduces reliability.

The method in [5] determines a region called region of embedding (ROE) and used this region to embed patient information and hash value of ROI for tamper detection. The ROE portion of the image must accommodate special blocks that ensure the rule given by the authors. The radiologist is responsible to determine ROE in a way that it does not intersect with ROI. The hash value of medical image computed by a hash function is used to detect tamper in this paper.

Difference expansion (DE) technique is proposed in [6] to embed watermark medical images with tamper localization capability. Their method determines a region called region of authentication (ROA) and divides it into subregions in a hierarchical manner. Signatures are produced from each subregion, and they are embedded into the medical image as payload. ROI shading is utilized to detect tampered regions in this work.

Two-dimensional DE (2D-DE) is used by Al-Quershi et al. in their work [7]. The method assumes that medical image consists of three parts: ROI, RONI, and border pixels. Hash value of ROI, patient information, pixels of ROI, and LSBs of border pixels are compressed with Huffman coding. Compressed stream is embedded into RONI using 2D-DE. However, the location of the map that is the result of DE and the coordinates of ROI are embedded into border pixels. The most important problem with this scheme is limited ROI size which can be up to 12% of the original image. Another disadvantage is localization of the tampered region since the whole ROI is designated as tampered even if one pixel in ROI is tampered.

Al-Quershi et al. used average pixel intensities of blocks in ROI to detect the tampered regions [8]. Average values from the blocks in ROI are gathered as payload before watermark embedding. The scheme embeds EPR into ROI using a reversible technique based on DE. Average values of blocks are embedded into RONI using a robust watermarking technique. Retrieved average values during watermark extraction are used to determine the tampered regions. The method uses lossy compressed ROI for recovery purposes. Using the average value for tamper localization is not appropriate if the modification on the block does not affect the average value.

The methods in [8, 9] are designed to provide the four properties: hiding the patients' data, authentication of ROI, tamper localization in ROI, and recovering the tampered regions. The JPEG 2000 compressed version of ROI is used as recovery information. While ROI carries patient data, tamper localization and recovery information are embedded into RONI. MDE proposed in [5] is used by both of these techniques. While the method in [9] used DE described in [1] to embed information into RONI, a three-level discrete wavelet transform (DWT) was used for data embedding in [8].

The researcher proposed a method based on DE [10]. The method divides the image into 16×16 pixel blocks and uses their cyclic redundancy check (CRC) values for tamper localization. Each block accommodates its own authentication value. The following blocks are used if the CRC value cannot be embedded into corresponding block due to the DE technique. Whole watermark is extracted to determine the tampered regions. Chaining mechanism during embedding will result in error propagation during watermark extraction algorithm.

The method proposed in [11] reverses the ROI of original medical image after watermark extraction. The LSB information of the ROI region has been embedded into RONI, and average values of 4×4 blocks in ROI have been used to recover the ROI if a tampering occurred on this region. However, one bit for authentication and one bit for parity check are used to locate the tampered regions. Only two authentication bits for 8×8 blocks are used for tamper detection.

Das et al. partition the image into two sections: ROI and RONI [12] in 2013. Two LSB bits of ROI, hash of ROI, and coordinates of ROI and EPR are concatenated and replaced by second LSBs of medical image after arithmetic coding. After that, the method divides the watermarked image into 3×3 pixel regions. Nine predetermined bits are obtained from hash value of each block, and each block's LSBs are modified to accommodate its own authentication bits. Bit planes after the second bit are not robust to any attack in this method.

Liew et al. proposed another region-based algorithm to watermark the medical images in 2013 [13]. Their method divides the medical image into ROI and RONI with non-overlapping blocks of size 40×40 and 2×2 , respectively. While one portion of RONI is used for hiding the authentication data, another portion accommodates the recovery data. Hash and CRC result of ROI blocks are embedded into RONI, and the JPEG 2000 compressed version of ROI is used as recovery information. This method designates a 40×40 pixel region as tampered even if one pixel of it is modified.

Huang proposed a fragile watermarking method to protect the integrity of medical images [14]. His method also detects tampered regions on the medical images. Their results show that the method can accurately localize the tampered regions. Two LSBs of the medical image is used for watermark embedding. However, the method is not reversible and the original medical image cannot be obtained after watermark extraction.

Eswaraiah et al. proposed a fragile block-based medical image watermarking technique [15]. Their technique divides the image into three sections (ROI, RONI, and border pixels) and provides four properties: no distortion inside ROI, integrity verification of ROI, tamper localization in ROI, and recovering the original ROI. While border pixels contain authentication and information data of ROI, recovery data of ROI is embedded into RONI. Two, 3, or 4 LSBs of RONI are modified to embed the corresponding ROI pixel values.

ROI authentication is realized with a hash function. Their method assumes the whole ROI region tampered even if one bit in ROI changes by mistake.

Al-Haj et al. proposed a region-based algorithm that utilizes multiple watermarking in the frequency and spatial domains in 2014 [16]. While robust watermarking is used for RONI, fragile watermarking with reversible capabilities is used for ROI. LSBs of the ROI section, EPR, and hospital's logo are constituted the watermark which is embedded into RONI using singular value decomposition (SVD)- and DWT-based robust watermarking technique. Whole LSBs of ROI cannot be embedded into the RONI region due to the size of ROI, and therefore, the algorithm randomly selects some ROI pixels LSBs to form the watermark. The algorithm embeds fragile watermarking pattern on those LSBs saved in RONI. Fragile watermark is pseudo-random noise (PN) sequences. Unfortunately, the method cannot detect tamper except randomly selected pixels.

Thabit et al. exploits two watermarking approaches based on slantlet transform (SLT) to embed data into ROI and RONI [17]. Their method generates recovery information using the integer wavelet transform (IWT) coefficients. ROI is selected by clinicians and then divided into non-overlapping 16×16 pixel blocks. Average pixel intensities and recovery information from ROI blocks are calculated using the IWT. While recovery information and average values are embedded into RONI using a robust irreversible technique, EPR information is embedded into ROI with a reversible technique. Two drawbacks of this method are that it uses average information from 16×16 blocks to detect the tampered regions, and it must send some side information with watermarked medical image.

In [18], Eswaraiah et al. proposed a medical image watermarking technique based on the IWT. The medical image is segmented into ROI and RONI regions. Hash of ROI, recovery information of ROI, and EPR are embedded into RONI using IWT. The disadvantages of the method are that the coordinates of ROI and the size of watermark are sent to the other side as side information, the authentication of ROI depends on hash function (one bit modification in ROI will necessitate the recovery of ROI), and it can be applied to only medical images whose ROI size do not exceed 20% of the whole image.

Some requirements of the medical image watermarking techniques can be given as the following:

- ROI size restriction
- Need for clinical assistance to select ROI
- Size of the marked region when one pixel has been modified
- Reversibility of the watermarking method
- Enhancing the watermarking capacity while preserving visual quality of the medical image

We aim to improve the requirements listed above with a new reversible medical image watermarking technique which uses the MDE watermarking technique to hide EPR information and authentication bits in this work. The method divides the center area of the medical image into non-overlapping 4×4 pixel blocks. Four 2×2 sub-blocks in each block are then evaluated as expandable or non-expandable using the MDE technique defined in the “MDE Technique” section. The number of expandable sub-blocks in the current block determines the strategy to be used for authentication bit generation and embedding for the current block. Generated location map during embedding phase and its hash value are also embedded into the border pixels of the medical image. The first advantage of the algorithm is that it does not necessitate dividing the image into regions (ROI, RONI) and uses whole pixels except the border pixels in the image to embed the watermark. Region-based methods in the literature fail to embed watermark for medical images (ultrasound, X-ray images, etc.) with larger ROI compared to RONI. Another problem with region-based methods is the determination of the exact borders of the ROI region. The second advantage of the proposed method is that it can detect tampered regions with finer precision. A small 4×4 pixel window is marked as tampered when one pixel is modified. The method’s marked pixels per modified pixel ratio is quite good compared to other methods in the literature. Experimental results show that the method can detect tampered regions with high accuracy (with smaller window size) and ROI size of the medical image does not affect the algorithm. The method generates watermarked images with approximately 43-, 38-, and 39-dB peak signal to noise ratio (PSNR) for X-ray, magnetic resonance imaging (MRI) and ultrasound (US) modalities.

The rest of the paper is organized as follows: the “MDE Technique” section outlines MDE technique. The proposed method is explained in the “Proposed Method” section, and experimental results are given in the “Experimental Results” section. Conclusions are drawn in the final section.

MDE Technique

Guo et al. defined the modified difference expansion technique in their work to overcome some problems of the Alattar’s scheme [2, 5]. Their work assumes that a quad is a vector denoted by $u = (u_0, u_1, u_2, u_3)$ and represents 2×2 adjacent pixel values. The method also assumes that all quads are scanned in a non-overlapping manner. Forward difference expansion transform denoted by f gets the quad as parameter

and returns a 1×4 vector denoted by $v = (v_0, v_1, v_2, v_3)$. The function f calculates the value of v using Eq. 1.

$$\begin{aligned} v_0 &= \frac{u_0 + u_1 + u_2 + u_3}{4} \\ v_1 &= u_1 - u_0 \\ v_2 &= u_2 - u_0 \\ v_3 &= u_3 - u_0 \end{aligned} \quad (1)$$

Inverse difference expansion transform f^{-1} is also applied on v to retrieve the original value of u as in Eq. 2.

$$\begin{aligned} u_0 &= v_0 - \frac{v_1 + v_2 + v_3}{4} \\ u_1 &= v_1 + u_0 \\ u_2 &= v_2 + u_0 \\ u_3 &= v_3 + u_0 \end{aligned} \quad (2)$$

MDE labels the quad $u = (u_0, u_1, u_2, u_3)$ to be expandable for all watermark bits $b_1, b_2, b_3 \in \{0, 1\}$ if v can be modified to generate $\tilde{v} = (v_0, \tilde{v}_1, \tilde{v}_2, \tilde{v}_3)$ using Eq. 3 without causing any overflow or underflow during inverse transform $f^{-1}(\tilde{v})$:

$$\begin{aligned} v_0 &= \frac{u_0 + u_1 + u_2 + u_3}{4} \\ \tilde{v}_1 &= 2v_1 + b_1 \\ \tilde{v}_2 &= 2v_2 + b_2 \\ \tilde{v}_3 &= 2v_3 + b_3 \end{aligned} \quad (3)$$

If the inverse transform $f^{-1}(\tilde{v})$ do not cause any overflow or underflow on the pixel values (below 0 or above 255 for a depth of 8 bits), the method labels the quad as expandable and embeds a 3-bit watermark information b_1, b_2, b_3 into the corresponding block. The method is reversible because the values of $\tilde{v}_1, \tilde{v}_2, \tilde{v}_3$ are the 1 bit left shifted version of the original values calculated in Eq. 1. Thus, the inverse transform given in Eq. 2 reverses the original pixel values in the quad after watermark bit extraction.

Proposed Method

The details of the proposed method are given in this section. The method consists of two parts: watermark embedding and watermark extraction algorithms. Watermark embedding algorithm embeds the authentication bits into corresponding blocks and EPR information into blocks, and extraction algorithm verifies the extracted authentication bits and finds the tampered regions. EPR information is also extracted during the watermark extraction algorithm.

Watermark Embedding Algorithm

The proposed method segments the medical image into two parts: border area (BA) and center image area (CIA). Figure 1

Fig. 1 Border area (BA) and center image area (CIA) of the medical images. **a** MRI. **b** Ultrasound (US). **c** X-Ray

shows these regions for three different medical image modalities. Pixels between the lines form the BA region and the rest of the image is the CIA region. The method does not need to segment medical image into ROI and RONI regions thus it does not necessitate any ROI selection algorithm or side information about the coordinates of ROI during the embedding algorithm. The whole medical image is processed by the method in the same way. Thus, the method can be applied to any image modalities such as computed tomography (CT) images that have the larger ROI region.

The algorithm can be summarized as follows: the method divides CIA into N non-overlapping 4×4 pixel main blocks $B_1 \cdots B_N$ and applies the *block classification* procedure to each block. This procedure tests the expandability of the current block and then classifies the block as expandable or non-expandable and also updates the location map (LM) for sub-blocks. After this procedure, the method calls the *block feature value (BFV) generation* procedure to compute BFV for both expandable and non-expandable blocks by separate algorithms. When all blocks are processed by these two

procedures, the *payload and EPR embedding* procedure is called by the method to embed patient record and necessary information for reversible capability.

The watermark embedding procedure shown in Fig. 2 is given below.

- Step 1. Segments the medical image into two areas: BA and CIA.
- Step 2. Divides CIA into N non-overlapping 4×4 sub-blocks $B_1 \cdots B_N$.
- Step 3. Labels each block $B_1 \cdots B_N$ using *block classification procedure* as expandable or non-expandable.
- Step 4. Applies *BFV generation and embedding procedure* to each block using its label.
- Step 5. Embeds the EPR information and payload into the blocks labeled as expandable using the *payload and EPR embedding procedure*.

The details of each procedure in the embedding phase are given in detail as the following.

Fig. 2 Flowchart of watermark embedding algorithm

Block Classification Procedure

The procedure divides current block B_i into four non-overlapping 2×2 sub-blocks X , Y , Z , and W . Then, the MDE technique as explained in the “MDE Technique” section is applied to these four sub-blocks to test and determine their expandability. For the first sub-block X , Eqs. 1 and 3 are used to determine whether it is expandable or not. Pixels in the sub-block are denoted by (x_1, x_2, x_3, x_4) . The block is marked as expandable if the calculated value

of $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4)$ for all values of $b_1, b_2, b_3 \in \{0, 1\}$ do not cause underflow or overflow. Otherwise, the block cannot be used for data embedding by MDE. $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4)$ denotes the pixel values for the current block after three watermark bits b_1, b_2, b_3 are embedded. Other sub-blocks Y, Z , and W are also tested for expandability likewise.

The procedure classifies current 4×4 block B_i as expandable if more than two of its sub-blocks are expandable. Otherwise, the block is labeled as non-expandable.

The procedure also inserts four binary values into the LM to represent the expandability of its sub-blocks. For example, if the first and last sub-blocks in the current block are expandable, binary value 1001 is inserted into the LM. Thus, a 4-bit information is inserted into LM to represent the expandability of the sub-blocks of each block.

BFV Generation and Embedding Procedure

The procedure extracts a 1×6 feature vector denoted by BFV_i from each block according to its label by using the steps defined below. Assume that pixel values in the i th block are denoted by $B_i^1 \cdots B_i^{16}$.

Step 1. If the label of the current block is expandable, the procedure computes the corresponding hash value H_i for the current block pixel values using well-known 256-bit secure hash algorithm, SHA-256.

If the label of the current block is non-expandable, the LSB of the pixels are set to zero and then SHA-256 is applied on the newly created block pixels $B_i^1 \cdots B_i^{16}$ to compute the hash value H_i .

Step 2. The predetermined set of pixels' $B_i^1, B_i^3, B_i^6, B_i^9, B_i^{11}, B_i^{16}$ values in the current block is used to index the current hash value H_i , and BFV_i for the current block is calculated as in Eq. 4. Assume that the z th bit value of H_i is denoted by $H_i(z)$.

$$BFV_i = H_i(B_i^1) \| H_i(B_i^3) \| H_i(B_i^6) \| H_i(B_i^9) \| H_i(B_i^{11}) \| H_i(B_i^{16}) \quad (4)$$

Step 3. Step 3.1. If the label of the current block is expandable, the 6-bit block feature value BFV_i is embedded into the first two expandable blocks using Eqs. 1 and 3.

Step 3.2. If the label of the current block is non-expandable, the LSBs of all the four pixels in X , (x_1, x_2, x_3, x_4) , and the first two pixels in Y , (y_1, y_2) , are kept in an array called RES1 and LSB embedding will be used to hide BFV_i into the current block B_i . These LSBs (six LSBs) are modified to hide the current block feature value BFV_i . Figure 3 also shows an example for BFV generation. The figure indicates various strategies by the procedure according to the label of the current block.

The MDE technique embeds 3-bit information into 2×2 expandable blocks as given in Step 3.1. If the current block is

not expandable, MDE cannot use the current block to embed watermark information. The proposed method uses the LSB embedding technique to increase the embedding capacity for these blocks. Thus, the method can embed more watermark information into the same medical image when compared to a method which uses only MDE. We do not use only LSB embedding because it modifies the medical image irreversibly, and there is no way to recover the original pixel values in the image. Figure 4 shows the superiority of the method. Assume that the current block has two expandable and two non-expandable sub-blocks denoted by 0110. Figure 4a indicates that a 6-bit watermark can be embedded into 4×4 blocks if only MDE is used. The proposed method uses both MDE and LSB and can embed 14 watermark information bits into the same block as shown in Fig. 4b.

Payload and EPR Embedding Procedure

After all the blocks were processed by the two procedures defined above, LM will be compressed using arithmetic encoding. The hash value of the original LM denoted by H , compressed LM, and its additional information is encrypted with a symmetric key encryption technique by using a key value K that will be shared between the sides. Crypted information is embedded into the pixels in BA using LSB embedding technique. The last two bits of pixels in BA will be kept in an array RES2. The array RES2 must be kept to recover the original values of pixels in BA during the watermark extraction procedure. The bit values in RES1, RES2, and EPR information are embedded into the blocks that are labeled as expandable.

However, the expandable number of blocks plays an important role during the embedding procedure. Assume that the number of non-expandable blocks denoted by num_{nonexp} and the number of expandable blocks that contain three and four expandable sub-blocks denoted by num_{exp3} and num_{exp4} , respectively. If the Eq. 5 is satisfied, the third and fourth sub-blocks in the expandable blocks are modified using Eqs. 1 and 3 to accommodate the bit values in $RES1 + RES2 + EPR$. Otherwise, the algorithm cannot embed the watermark data into the current medical image. $size(x)$ given in Eq. 5 returns the number of bits.

$$6 \cdot num_{nonexp} + size(RES2) + size(EPR) \leq 3 \cdot num_{exp3} + 6 \cdot num_{exp4} \quad (5)$$

3.2. Watermark Extraction Algorithm

In this phase of the method, the watermark extraction algorithm determines the tampered regions and

Fig. 3 Example of block feature value (BFV) generation

extracts EPR information from the medical image. The details of the method are given below, and Fig. 5

Fig. 4 Example of **a** Watermark Embedding with only MDE and **b** watermark embedding with MDE + LSB

shows the flowchart of the watermark extraction algorithm.

Encrypted information (hash value of original LM denoted by H , compressed LM, and its additional information) extracted from BA is decrypted using shared symmetric key between the sides. Extracted LM is decompressed to lead the watermark extraction algorithm. The algorithm should extract the LM first to determine the MDE-modified 4×4 blocks in CIA. The rest of the CIA blocks are modified by LSB embedding. The hash value is again computed on the extracted LM' and compared with the extracted value H' . Watermark extraction algorithm uses LM' to determine embedding type if both hash values are equal. CIA is again divided into N 4×4 blocks $B_1 \dots B_N$. All blocks in CIA are scanned to extract $RES1' + RES2' + EPR'$ information. The third and fourth 2×2 sub-blocks in 4×4 blocks with more than two expandable blocks are used to extract $RES1' + RES2' + EPR'$ information.

Fig. 5 Flowchart of watermark extraction algorithm

MDE technique is used on these blocks to extract information and recover the original pixel values in these blocks. After all the blocks are processed, $RES2'$ information is used to recover the original values of the pixels in the BA. EPR' is used to authenticate the patient if no tamper detection is found, and $RES1'$ is used to recover the original pixel values of pixels modified by LSB embedding to hide the BFV value in CIA during the *tamper detection phase*.

At this point, LM' is extracted and verified, the original values of pixels in BA are recovered, EPR' information is extracted, and $RES1'$ information is extracted by the algorithm. Then, the algorithm computes the $BFVs$ of the blocks $B_1 \cdots B_N$ again and compares them with the extracted ones for *tamper detection phase* as the last step. The steps for this phase are given below.

Step 1. Divide CIA into N non-overlapping 4×4 main blocks, $B_1 \cdots B_N$.

Step 2. Repeat the following steps for $i = 1 \cdots N$

Step 2.1. Divide the current block B_i into four 2×2 sub-blocks X , Y , Z and W .

Step 2.2. Extract four bits from LM' and determine how many sub-blocks were marked expandable from these bits.

If two or more sub-blocks are expandable,

- Compute block feature value BFV_i for the current block using the *BFV generation procedure*.
- Extract embedded BFV'_i from the first two sub-blocks using MDE and compare it with the computed BFV_i .
- If they are equal, reconstruct original pixel values in the first two sub-blocks using MDE. Otherwise, mark the corresponding block B_i as tampered.

Else,

- Extract the embedded BFV'_i which is extracted from LSBs from four pixels in the first sub-block and one LSB from the first two pixels in the second sub-block and compare it with BFV_i .
- If they are equal, recover the original LSBs from $RES1'$. Otherwise mark the corresponding block B_i as tampered.

Experimental Results

Experimental results of the proposed method are summarized for imperceptibility; tamper localization and embedding capacity in this section. Scale medical images (210 Gy) of three modalities (MRI, US, and X-ray) of size 268×268 , 512×512 , 1024×1024 , and 4096×4096 with 1 K EPR information have been used during the experiments. The method is coded in MATLAB R2014b on a notebook computer with Intel Core i7 5500U 2.4 GHz CPU, and 8 GB RAM. Three example test images from the dataset are shown in Fig. 1. While the area between the yellow lines shows the region of BA, inner pixels will constitute the region of CIA. Table 1 shows the size of the used test images, CIA, and BA regions.

The first experiment in this section shows the imperceptibility of the proposed method. PSNR is used to measure the quality of the watermarked medical image. The equation given in Eq. 6 determines the PSNR between the original medical image M and the watermarked medical image W of size $X \times Y$.

$$\text{PSNR} = 10 \log \left(\frac{255^2}{\text{MSE}} \right) \quad (6)$$

$$\text{MSE} = \frac{\sum_{i=1}^X \sum_{j=1}^Y (|M_{ij} - W_{ij}|)}{X Y}$$

Figure 6a, c, e show the original medical images and Fig. 6b, d, f) are the corresponding watermarked medical images. PSNRs for MR, US, and X-Ray images are approximately 36, 37, and 44 dB respectively. The X-ray image has higher PSNR because it has more expandable blocks compared to the other two medical images. The number of expandable blocks affects the visual quality of the watermarked medical image. However, changes on the watermarked medical images given in Fig. 6b, d also do not draw attention of human eyes.

The method is tested on 210 tampered medical images, and the average PSNR is calculated. Table 2 shows the average PSNRs of the proposed method for different modalities. While PSNR for MRI and US images is nearly the same, X-ray images have higher PSNR because they have more expandable blocks compared to others. We also compare the proposed method with similar works in the literature to test the effectiveness of the method according to invisibility. Figure 7 shows the average PSNR values for the proposed method and similar methods [16, 17] with different medical image modalities of size 2048×2048 . While the proposed method has approximately 47 dB PSNR, the others have approximately 45 and 34 dB PSNR, respectively [16, 17]. When MRI and US test images are considered, the proposed method also has better PSNR compared to similar methods [16, 17].

The tamper localization capability of the proposed method is also tested in another experiment. This experiment shows the capability of the method for three cases. In the first case, only one pixel is modified on the test image and the tamper localization capability of the method is tested. The second case considers the situation where a part from the medical image is copied and pasted onto another region on the same medical image. Copied and pasted regions have regular perimeter. The third case evaluates the method when copied and the pasted regions have irregular perimeter.

Figures 8 and 9 show the results of the proposed method for the first case. Two distinct one-pixel values in two test images given in Fig. 8a, d are modified, and tampered medical images are obtained as in Fig. 8b, e. Red circles in the tampered

Table 1 Size of the test images, CIA, and BA

Modality	Size of image (pixels)	Size of CIA (pixels)	Size of BA (pixels)	Location map size
X-ray	268×268	65,536	6288	16,384
	512×512	238,144	24,000	59,536
	1024×1024	976,144	72,432	244,036
	4096×4096	16,289,296	487,920	4,072,324
MRI	268×268	65,536	6288	16,384
	512×512	238,144	24,000	59,536
	1024×1024	976,144	72,432	244,036
	4096×4096	16,289,296	487,920	4,072,324
US	268×268	65,536	6288	16,384
	512×512	238,144	24,000	59,536
	1024×1024	976,144	72,432	244,036
	4096×4096	16,289,296	487,920	4,072,324

Fig. 6 **a** MRI, **c** ultrasound, and **e** X-Ray—original medical images of size 268×268 . **b** MRI, **d** ultrasound, and **f** X-ray—watermarked medical images of size 268×268

images enclose the modified pixels. The proposed algorithm detects the tampered 4×4 regions as shown in Fig. 8c, f. Tagged regions contain modified pixel values.

Tampered images given in Fig. 9b, e are created to show the effectiveness of the method for the second and third cases. While Fig. 9b contains tampered areas with irregular shape, regular areas are used in Fig. 9e. The proposed method can detect tampered regions for both cases as shown in Fig. 9c, f.

The method is also evaluated according to its embedding capacity. Table 3 shows the average embedding capacity of BA and CIA regions, and also, the average total embedding

capacity of the proposed method is indicated in it. For example, the average total embedding capacity for MRI images of

Table 2 Average PSNR for three modalities

Modality	X-ray	MRI	US
268×268 images	41.48	36.29	36.80
512×512 images	43.81	38.21	38.51
1024×1024 images	45.83	42.17	42.13
4096×4096 images	49.54	46.83	47.13

Fig. 7 PSNR comparison of the proposed method with [16] and [17]

size 268×268 in the test dataset is approximately 54,274 bits. There is only 54,148 of them used for embedding the watermark and EPR information. The last column in the table indicates the average total embedded BFV for various medical image modalities.

We also evaluate the proposed method according to invisibility and capacity metrics. EPR information (1 K, 4 K, 16 K, and 64 K) are used in the experiment to test the invisibility for various sizes of EPR information. Three different image

modalities are used during the experiment. Figure 10 indicates that PSNR decreases as EPR increase. Average PSNR for X-ray, MRI, and US modalities is computed as 44.61, 41.25, and 41.57 dB respectively, for 64 K EPR information embedded into the test medical images. We choose 1 K EPR because it gives enough information and invisibility as shown in the experiment.

Another experiment is realized to test the detection performance of the proposed method. Tampering detection performance is measured using detection sensitivity or true positive rate (TPR). TPR is a measure to identify positive results. It also determines the probability of correctly identifying tampered blocks. The equation used for calculation of TPR is given in Eq. 7 where true positive blocks (TP) indicates the number of tampered blocks correctly identified as tampered and false negative blocks (FN) designates the number of tampered blocks incorrectly identified as unmodified:

$$TPR = \frac{TP}{TP + FN} \times 100 \quad (7)$$

Table 4 lists the average TPR results for the method for three different modalities. In the experiment, tampered regions

Fig. 8 **a** Ultrasound and **d** X-ray, watermarked images. **b** Ultrasound and **e** X-ray, two pixel-modified watermarked images. **c** Ultrasound and **f** X-Ray, tamper-localized images

Fig. 9 **a** MRI and **d** ultrasound, watermarked images. **b** MRI and **e** ultrasound, region modified watermarked images. **c** MRI and **f** ultrasound, tamper-localized images

with various sizes are used. Size of the tampered region is determined by a ratio denoted by p , which denotes the ratio of the size of the tampered region to the size of the original image. For example, the size of the modified region becomes

32×32 when p is chosen as 1.43% for a 268×268 MR image. The experiment shows that the algorithm can detect approximately 99% of tampered regions. Another metric used for tamper detection performance is false positive rate (FPR) that

Table 3 Embedding capacity of the proposed method

Modality	Image size	Avg. BA embedding capacity (bits)	Avg. CIA embedding capacity (bits)	Avg. total embedding capacity(bits)	Avg. total embedded data(bits)	Avg. total embedded BFV (bits)
X-ray	268×268	12,576	41,458	54,034	51,175	24,576
	512×512	48,000	151,360	199,360	189,550	89,304
	1024×1024	144,864	644,679	789,543	662,567	366,054
	4096×4096	975,840	12,153,361	13,129,201	12,302,439	6,108,486
MRI	268×268	12,576	41,698	54,274	54,148	24,576
	512×512	48,000	150,365	198,365	198,312	89,304
	1024×1024	144,864	664,073	808,937	685,907	366,054
	4096×4096	975,840	12,958,457	13,934,297	13,165,833	6,108,486
US	268×268	12,576	40,783	53,359	53,323	24,576
	512×512	48,000	149,198	197,198	196,660	89,304
	1024×1024	144,864	653,581	780,445	672,308	366,054
	4096×4096	975,840	12,458,873	13,434,713	12,655,765	6,108,486

measures the ratio of incorrectly identifying unmodified blocks as tampered. However, the algorithm will extract exact watermark information from the medical image if it is not modified. Therefore, the FPR value has been calculated to be zero during the tests. A malicious user can modify a block with $1/64$ (since each block accommodates only 6-bit BFV) probability in an undetectable manner. If the proposed method chooses smaller block size such as 2×2 , each block can accommodate a maximum of 3-bit BFV. It increases the probability of a malicious user to create a tampered block in an undetectable manner and also decreases TPR of the proposed method.

The experiments compare the proposed method with similar methods in the literature to make a fair comparison. The methods are listed with their capabilities. Table 5 shows that the proposed method detects a tampered pixel with a 4×4 window. Other methods except [12] and [18] mark the tampered regions with larger windows. However, the method in [12] has ROI size restriction. The method encounters problems for medical images with larger ROI. In the same manner, the method in [18] has also ROI size restriction and it only reverses the ROI region. The proposed method marks the tampered region accurately by 4×4 windows and it does not have a ROI restriction. Also, reversibility is not bounded only to the ROI region for the method.

It has been observed that some medical images (especially 268×268 MRI images) contain more textured areas compared to smooth areas. Thus, the proposed method cannot embed watermark information into this type of medical images. Two percent of MRI images of the current dataset cannot be watermarked for this reason with the proposed method.

Discussion and Concluding Remarks

In this paper, we used a hybrid (MDE + LSB) region-based watermarking technique to verify the authenticity and to ensure the integrity of medical images with different modalities.

Fig. 10 Average PSNR values for the different EPR sizes

Table 4 Average TPR of the proposed method

Modality	$p = 1.43\%$	$p = 2.85\%$	$p = 4.28\%$	$p = 5.70\%$
X-ray (268×268)	98.98	98.78	98.72	99.04
MRI (268×268)	99.06	98.71	98.67	98.82
US (268×268)	99.29	98.79	98.82	98.77

The proposed method is reversible and has increased watermarking capacity. The method does not necessitate dividing the image into ROI and RONI. Thus, it does not necessitate a clinician or an algorithm to decide ROI before the embedding procedure. Each 4×4 block accommodates 6-bit BFV generated from the current block. Either MDE or LSB embedding method is used on blocks to embed BFV depending on their expandability characteristics. LSB embedding operation generates a payload which is embedded into the BA of the medical image after being concatenated with 1 K EPR. The performance of the method has been evaluated with respect to PSNR, embedding capacity, and TPR for various medical image modalities such as X-ray, MRI, and US by the experimental results. Tables 2, 3, and 4 give the effectiveness of the proposed method with different aspects.

In this work, we propose a new medical image watermarking method that uses both the MDE and LSB embedding techniques. The method divides the image into two regions: BA and CIA. Authentication bits for each block in CIA are created and embedded into corresponding blocks using MDE. The LSB embedding technique is used to embed authentication bits if a 4×4 block has not enough expandable 2×2 sub-blocks. BA accommodates both LM and its hash value. Thus, tampered regions can be detected accurately with smaller window size and ROI size does not affect the watermarking algorithm. Experimental results indicate that the method generates approximately 43-, 38-, and 39-dB PSNR watermarked images for X-ray, MRI, and US modalities. Table 4 also shows that the method detects tampered regions with smaller window size, and it has no ROI size restriction. It is planned to improve the watermarking algorithm and apply the method to thermal medical images.

The proposed method aims to improve the detection accuracy by using 4×4 blocks. Thus, the size of the location map becomes larger to provide finer localization of the tampered area. It is planned to either remove or decrease location map size. We will investigate new reversible watermarking methods to improve the embedding capacity of the proposed method in the future. We also plan to improve the interoperability of the method by enriching the watermark information with DICOM header. Utilizing DICOM header for watermark embedding and transferring medical image in DICOM format would certainly provide interoperability between the PACS which has general use in the clinical world.

Table 5 Comparison of the methods with similar works according to basic properties

Methods	Objectives	ROI-based	Embedded TECHNIQUE	Embedded DATA	Tamper LOCALIZATION	Tamper DETECTION SIZE	Applicable to medical images of any size and modality	Reversibility	Accurate identification of tampered blocks	ROI size restriction
[3]	Authentication, data hiding	X	LSB + DWT	Authentication and patient's data	X	–	✓	X	X	X
[4]	Authentication	X	Frequency domain DCT	Authentication and recovery data	✓	–	✓	✓	✓	✓
[5]	Authentication, data hiding	✓	MDE	Authentication and patient's data	X	–	✓	✓	X	X
[6]	Authentication, data hiding	✓	MDE	Authentication and patient's data	✓	$80 \times 60, 32 \times 40$	✓	✓	✓	X
[7]	Authentication, data hiding	✓	DE	Authentication, recovery and patient's data	✓	16×16	X	✓	✓	✓
[8]	Authentication, data hiding	✓	MDE + DWT	Authentication, recovery and patient's data	✓	16×16	X	Only ROI	X	✓
[9]	Authentication, data hiding	✓	MDE + DE	Authentication, recovery and patient's data	✓	16×16	✓	Only ROI	✓	✓
[11]	Authentication, data hiding	✓	Spatial domain LSB	Authentication, recovery and patient's data	✓	8×8	X	Only ROI	X	✓
[12]	Authentication, data hiding	✓	Spatial domain LSB	Authentication and patient's data	✓	3×3	✓	✓	✓	✓
[13]	Authentication, data hiding	✓	Spatial domain LSB	Authentication, recovery and patient's data	✓	40×40	X	Only ROI	✓	✓
[16]	Authentication, Data Hiding	✓	LSB + DWT + SVD	Authentication and patient's data	✓	2×2	✓	Only ROI	✓	✓
[17]	Authentication, Data Hiding	✓	Frequency domain SLT	Authentication, recovery and patient's data	✓	16×16	✓	Only ROI	✓	✓
[18]	Authentication, data hiding	✓	Frequency domain IWT	Authentication, recovery and patient's data	✓	3×3	✓	Only ROI	✓	✓
Our Method	Authentication, data hiding	X	MDE + LSB	Authentication and patient's data	✓	4×4	✓	✓	✓	X

References

1. Tian, J., "Reversible watermarking using a difference expansion", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, 2003
2. Alattar, A. M., "Reversible watermark using the difference expansion of a generalized integer transform", *IEEE Trans. Image Process.*, vol. 13, no. 8, pp. 1147–1156, 2004
3. Woo, C-S, Du, J, Pham, B, "Multiple Watermark Method for Privacy Control and Tamper Detection in Medical Images". In *Proceedings of the APRS Workshop on Digital Image Computing Pattern Recognition and Imaging for Medical Applications*. 2005
4. Wu, J. H. K., Chang, R.-F., Chen, C.-J., Wang, C.-L., Kuo, T.-H., Moon, W. K., & Chen, D.-R. Tamper detection and recovery for medical images using near-lossless information hiding technique. *Journal of Digital Imaging*, 21(1), 59–76, (2008).
5. Guo, X., Zhuang, T. . A region-based lossless watermarking scheme for enhancing security of medical data. *Journal of Digital Imaging: The Official Journal of the Society for Computer Applications in Radiology*, 22(1), 53–64, (2009).
6. Guo, X., Zhuang, T. . Lossless watermarking for verifying the integrity of medical images with tamper localization. *Journal of Digital Imaging: The Official Journal of the Society for Computer Applications in Radiology*, 22(6), 620–628, (2009)
7. Al-Qershi, O. M., Khoo, B. E., "ROI-based tamper detection and recovery for medical images using reversible watermarking technique," *Information Theory and Information Security (ICITIS)*, 2010 I.E. International Conference on Beijing, 2010, pp. 151–155.
8. Al-Qershi, O. M., & Khoo, B. E. Authentication and data hiding using a hybrid ROI-based watermarking scheme for DICOM images. *Journal of Digital Imaging*, 24(1), 114–125, (2011). doi: [10.1007/s10278-009-9253-1](https://doi.org/10.1007/s10278-009-9253-1)
9. Al-Qershi, O. M., & Khoo, B. E., Authentication and data hiding using a reversible ROI-based watermarking scheme for DICOM images, *Proceedings of International Conference on Medical Systems Engineering (ICMSE 2009)* (2009), pp. 829–834
10. Tan, C. K., Ng, J. C., Xu, X., Poh, C. L., Guan, Y. L., & Sheah, K. . Security protection of DICOM medical images using dual-layer reversible watermarking with tamper detection capability. *Journal of Digital Imaging*, 24(3), 528–540, (2011).
11. Kulkarni, M. B., Patil, R. T., Tamper Detection & Recovery in medical image with secure data hiding using reversible watermarking, *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 3, pp. 370–373, 2012.
12. Sudeb Das and Malay Kumar Kundu, Effective management of medical information through ROI-lossless fragile image watermarking technique. *Comput. Methods Prog. Biomed.* 111, 3 , 662–675, (2013).
13. Liew S, Way S, Zain J, Tamper localization and lossless recovery watermarking scheme with ROI segmentation and multilevel authentication. *J Digit Imaging* 24:114–125, 2012
14. Liangyong, Huang A fragile watermark method for improving medical images security: *Journal of Computers*, Vol. 9, No. 9, pp 2043–2049, (2014), Academy Publisher.
15. R. Eswaraiah and E. Sreenivasa Reddy, "Medical image watermarking technique for accurate tamper detection in ROI and exact recovery of ROI" *International Journal of Telemedicine and Applications*, vol. 2014, 2014. doi:[10.1155/2014/984646](https://doi.org/10.1155/2014/984646), Article ID 984646, 10 pages
16. Al-Haj A. Secured telemedicine using region-based watermarking with tamper localization. *J Digit Imaging*.;27(6):737–50, 2014.
17. Thabit R, Khoo BE. Medical image authentication using SLT and IWT schemes. *Multimed Tools App.*, ;1–24, 2015 [10.1007/s11042-015-3055-x](https://doi.org/10.1007/s11042-015-3055-x).
18. 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," in *IET Image Processing*, , vol. 9, no. 8, pp. 615–625, 2015.