



# ROI-based reversible watermarking scheme for ensuring the integrity and authenticity of DICOM MR images

Asaad F. Qasim<sup>1,2</sup>  · Rob Aspin<sup>1</sup> · Farid Meziane<sup>1</sup> · Peter Hogg<sup>3</sup>

Received: 26 March 2018 / Revised: 28 November 2018 / Accepted: 3 December 2018

Published online: 15 December 2018

© The Author(s) 2018

## Abstract

Reversible and imperceptible watermarking is recognized as a robust approach to confirm the integrity and authenticity of medical images and to verify that alterations can be detected and tracked back. In this paper, a novel blind reversible watermarking approach is presented to detect intentional and unintentional changes within brain Magnetic Resonance (MR) images. The scheme segments images into two parts; the Region of Interest (ROI) and the Region of Non Interest (RONI). Watermark data is encoded into the ROI using reversible watermarking based on the Difference Expansion (DE) technique. Experimental results show that the proposed method, whilst fully reversible, can also realize a watermarked image with low degradation for reasonable and controllable embedding capacity. This is fulfilled by concealing the data into ‘smooth’ regions inside the ROI and through the elimination of the large location map required for extracting the watermark and retrieving the original image. Our scheme delivers highly imperceptible watermarked images, at 92.18–99.94 dB Peak Signal to Noise Ratio (PSNR) evaluated through implementing a clinical trial based on relative Visual Grading Analysis (relative VGA). This trial defines the level of modification that can be applied to medical images without perceptual distortion. This compares favorably to outcomes reported under current state-of-art techniques. Integrity and authenticity of medical images are also ensured through detecting subsequent changes enacted on the watermarked images. This enhanced security measure, therefore, enables the detection of image manipulations, by an imperceptible approach, that may establish increased trust in the digital medical workflow.

**Keywords** Medical imaging · Reversible watermarking · Difference expansion · DICOM · Integrity · Authenticity

## 1 Introduction

In most medical imaging domains traditional diagnosis has mostly migrated to e-diagnosis workflows. Hospital Information Systems (HIS) and medical imaging platforms generate and

---

✉ Asaad F. Qasim  
A.Qasim@edu.salford.ac.uk

manage digital images across many modalities including X-ray, Ultrasound, Magnetic Resonance (MR), Computerized Tomography (CT), etc. Images taken in a hospital are saved in the Picture Archiving and Communication Systems (PACS), and are typically managed within a digital workflow based on the Digital Imaging and Communications in Medicine (DICOM) standard [26].

The transmission of these medical images through, and across, hospitals, locations and administrative organizations, has become a common practice for many purposes within the digital medical workflow. These include diagnosis, treatment, training, distance learning and medical consultation between clinicians and radiologists [32]. In most cases, this is within the defined workflows of the PACS systems, but there are also many cases, both valid and occasionally nefarious, in which images and data are withdrawn from one system to be transferred to other institutions or people. During the process of production and exchange, medical images can be intentionally, or unintentionally, tampered with. This potentially has serious implications on the diagnosis of patients with possible life affecting impact outcomes, mortality, etc. Therefore, the ability to maintain the integrity and authenticity of these images has become significant, both within the internal systems and during their transfer to other systems [38].

Two methods are generally applied to ensure the integrity and authenticity of the image data: metadata and digital watermarking. In medical imaging, metadata includes the patient information connecting the image to the patient and medical report [24]. The most common metadata structure fulfills part 15 of the DICOM standard, where the data is saved in the image header [37]. This data includes information to describe the patient, image, and acquisition properties (part of the image file). Existing metadata techniques do not provide a secure relationship between the metadata and medical image. It is, therefore, easy to destroy, modify, or otherwise disconnect the metadata rendering the image unreliable. Digital watermarking is recognized as a robust approach to tackle these failings. However, current approaches have little consideration on specific aspects of medical imagery against the defined need for imperceptibility, reversibility, integrity control, and authentication [32, 38].

Digital watermarking is the hiding of data within the digital object. This data can later be detected/extracted to confirm the validity of the object [5]. In medical domains, if an image is modified during the workflow process a collapse in trust regarding the validity of the images is formed. Potentially, any small change to the image could lead to misdiagnosis or uncertainty, with possible life-threatening consequences or legal implications. Consequently, retrieval of the original data from the modified image is essential. Reversible or lossless watermarking approaches fulfill this requirement in that they guarantee the extraction of the watermark and full retrieval of the unmodified original image [27]. Medical image watermarking approaches can be classified into three schools: conventional methods, Region of Interest (ROI), Region of Non Interest (RONI), and reversible approaches [38]. In this research, we adopted a reversible watermarking technique which can completely retrieve the original unmodified image.

In conventional watermarking methods, the watermark is encoded in the entire cover image by replacing some of its details, typically Least Significant Bits (LSBs), or degrading some detail when using lossy compression methods [29]. Irreversible approaches are not suitable as they are not accepted by radiologists with unmodified images being favored for medical investigations [44]. Medical images have two parts; ROI and RONI. ROI region comprises the informative part of the image which is utilized for diagnosis and must be conserved without any degradation. However, RONI includes the non-critical part of the image (e.g. background). Occasionally this region may contain grey

level parts of slight interest [35]. Using the ROI part for hiding the watermark may deform the pixel intensities in this section which may lead to misinterpretations and consequently misdiagnosis. RONI watermarking techniques embed data in regions that are considered unimportant in medical examination. However, this can only be performed if RONI exists. The amount of data, that can be embedded, highly depends on the RONI size and ROI may not be preserved against malignant operations [38].

In medical applications, there are typically strict restrictions on data reliability that preclude any modifications, such as watermarking, that have a perceptible visual impact. Modifying a patients' medical image could affect their life by causing errors in reading and interpretations, which may lead to incorrect diagnosis and treatment. Therefore, fully reversible watermarking techniques, which can completely recover both the original unmodified image and the embedded watermark, are developed. Reversible watermarking approaches can be categorized into four groups: compression based [7, 10], histogram modification based [16, 21, 33], quantization based [22, 23] and Difference Expansion (DE) based [20, 28]. Reversible watermarking based on the DE technique are recommended by many recent studies, and typically exceed alternate reversible methods in terms of higher payload capacity and lower complexity [20, 28, 41].

Difference Expansion presented a new approach for reversible watermarking [45]. It encodes 1-bit into the LSB of the difference value of 2-pixels. Selected pixel pairs can either be any two adjacent pixels (horizontal or vertical) or any pairs of pixels selected in a pre-defined form. To raise the embedding capacity, Alattar [2] developed previous DE techniques by encoding 2-bits into the difference values of a 3-pixels. This extended technique utilizes spatial and spectral triplets of pixels to hide a pair of bits. Spatial triplets denote any 3-pixels selected from the identical spectral or color part of the image. Spectral triplets can also be any 3-pixels chosen from different spectral components. Further enhancement of hiding capacity is achieved by encoding 3-bits in the difference values of 4-pixels [3]. The easiest way of determining the pixel quad is by selecting consecutive  $2 \times 2$  pixel, and this approach can be further generalized [4]. This generalized method encodes several bits in the difference values vectors of connected pixels instead of pairs, triplets and quads. The weakness of the DE watermarking technique is the reduction of the hiding capacity due to the need for a location map denoting the pixels where data is embedded. The need for this map makes it difficult to control and predict the hiding capacity and distortion level because of the size of the location map [38].

We developed a novel reversible watermarking, based on the DE technique, for confirming authenticity and integrity of medical images which can be used to detect both intentional and unintentional manipulations. The proposed technique automatically segments the image into two parts: ROI and RONI, with the watermark encoded into smooth blocks ( $3 \times 3$  pixels) inside ROI. The main contributions of this research are:

- Hiding of the watermark in smooth regions inside the ROI part of the image. Smooth regions are defined as blocks that have least differences between their pixels values. This makes the deformation less visually perceptible.
- Evaluation of image distortion through clinical trials, based on relative Visual Grading Analysis (VGA). This enable identification of the level of modification that can be applied to medical images before modification is visually perceptible.
- Retrieval without location mapping. This significantly enhances hiding capacity whilst also reducing potential image degradation.

## 2 Related work

A range of methods can be used to verify the authenticity of digital medical images: [38].

- Hiding the Electronic Patient Record (EPR) to confirm that the information belongs to the correct patient.
- Hiding the metadata which is located in the header of DICOM images. However, because some metadata may be modified each time the image is distributed; only information related to the patient and image should be employed.
- Combining the header with the raw image data by concealing the Digital Signature (DS) of the header. Although this method decreases the message length, the header data is inextricably connected to the image during transmission.

Image integrity verification can typically be achieved by:

- Hiding the Digital Signature (DS) of the image.
- Hiding the Message Authentication Code (MAC) of the image.

At extraction, the integrity of the image can be validated by matching the recalculated DS/MAC and the previously hidden DS/MAC and identifying differences, if any, to determine applied modifications. In digital watermarking, there is an inverse relationship between the capacity, robustness, and imperceptibility. Therefore, an evaluated trade-off of properties may be applied depending on the desired application. The priority order of authentication and integrity applications is imperceptibility, robustness, and capacity [38].

In healthcare applications, a reversible, fragile and blind watermarking method is required for validating authenticity and integrity of medical images. Al-Qershi and Khoo [6] present two reversible watermarking approaches based on DE. The first approach combines a technique, which embeds 2 bits of the payload in each pair of pixels [45], with a scheme, which encodes 12 bits of the watermark into each smooth blocks of  $4 \times 4$  pixels [11]. The second method combines a technique, which embeds 3 bits of the watermark in each quad of pixels [3], with the same scheme in the first approach [11]. One of the special features of medical images, in comparison to nonmedical images, is the large 'smooth' areas (blocks with equal pixel values). These proposed approaches segment the image into smooth and non-smooth regions instead of ROI and RONI. High hiding capacity techniques are utilized in the smooth regions. However, DE is applied to the non-smooth regions. Although the scheme achieves high capacity, the major drawback is the lack of capacity control due to the need for embedding the compressed location map which is required for extraction.

Das and Kundu [14] developed a blind, fragile and ROI reversible watermarking scheme. This system joins lossless compression and encryption to hide DICOM metadata, image hash and tamper localization information into digital medical image. Secure Hash Algorithm (SHA-256) is adopted to calculate the ROI hash. Medical image integrity is confirmed by comparing the embedded and recalculated hash data. Eswaraiah and Reddy [15] presented a fragile watermarking method for validating the integrity of ROI, identifying the manipulated blocks inside ROI and recovering the original ROI region. In this technique, the medical image is divided into three parts; ROI, RONI and border region. The hash code of the ROI is computed using SHA-256 and hidden in the border region. Authentication and ROI recovery information are encoded into the RONI. Several limitations can be observed in these schemes [14, 15]; the

ROI part needs to be defined manually, the ROI can only be retrieved after extracting the concealed data, and a substantive location map is required for extraction.

A newer approach combines the features of reversible, zero and RONI watermarking methods [41]. This technique merges the zero-watermarking principle in the ROI using Dual Tree Complex Wavelet Transform (DT-CWT), with high capacity of reversible watermarking in the RONI. This scheme needs a location map to retrieve the embedded data and original unmodified image. Selvam, et al. [42] present a blind hybrid reversible watermarking approach, operating in transform domain, for increasing hiding capacity and protecting the medical image. Integer Wavelet Transform (IWT) and Discrete Gould transform (DGT) are used to encode the watermark within the medical image. In the extraction, the concealed watermark is retrieved, and the original unmodified image is restored without any auxiliary data. However, this approach exhibits high distortion with a low payload capacity.

Parah, et al. [36] propose a high capacity reversible watermarking system for content authentication of medical images. A Pixel to Block (PTB) conversion method was applied, to the cover image, to achieve high embedding capacity and confirm reversibility. The watermark, which consists of EPR, block checksum and logo bits, was encoded into the Intermediate Significant Bit (ISB) of the whole image to avoid LSB removal attacks. Although this scheme achieves high embedding capacity, the distortion of watermarked images is high. Gao, et al. [16] present a reversible watermarking approach to achieve contrast enhancement of ROI and tamper detection against attacks on the ROI. This scheme segments background and ROI of medical image using Otsu's thresholding method. The watermark is embedded along with distortion-less contrast enhancement of the ROI by expanding of the peak-pairs of the ROI histogram. The feature-bit matrix created from the ROI is encoded into the LSBs of the background pixels to ensure the reversibility of the ROI. The major limitations of this approach are the need for embedding the feature-bit matrix and only the ROI part of the image can be retrieved at extraction. A novel medical image authentication approach was proposed by Balasamy and Ramakrishnan [9] using Discrete Wavelet Transform (DWT) and Particle Swarm Optimization (PSO). This approach finds optimal wavelet coefficients for concealing the watermark data using PSO to produce a watermarked image with low distortion. The extraction process does not require auxiliary information, but this approach has high image deformation in comparison to the low hiding capacity of the technique.

Yang, et al. [46] propose a reversible and high capacity data hiding scheme for secure archiving of medical images. This scheme automatically segments the image into two parts; ROI and RONI. The contrast of the ROI part is enhanced by extending the grayscale values and encoding the data into peak bins of the extended histogram without stretching the histogram bins. The remaining large data is embedded into the RONI part without considering visual image quality. Evaluation of the scheme shows low invisibility between the original image and watermarked versions in terms of Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) due to applying contrast enhancement to the watermarked images. Pan, et al. [34] present a fragile reversible watermarking approach for digital radiographic images. This technique differentiates the background from anatomical details within the image. Histogram Shifting (HS) modulation is used to encode the watermark into the background section while HS is applied to wavelet detail coefficients of the anatomical object, encoding watermark data with the image quantum noise. This scheme delivers a reasonable visual image quality, but hiding capacity is very low. Atta-ur-Rahman, et al. [8] propose a blind reversible watermarking to realize high level of secrecy and integrity for medical images. This scheme utilizes a chaotic key to choose some pixels from the cover image to hide a chaotically created

watermark. The remainder of the pixels are transformed into residues by employing Residue Number System (RNS). A primitive polynomial, of degree four, is applied to divide the selected pixels and obtain the remainder which is appended to the watermark message. The validity of the watermark is ensured, at extraction, based on the calculated remainder. This approach exhibits high levels of imperceptibility, however, the embedding capacity of the scheme has not been measured. Moreover, the scheme does not rely on a region based watermarking strategy which makes the technique incapable of selecting the hiding regions.

Our research proposes a blind, fragile and reversible watermarking technique for encoding the DICOM metadata and DS of the whole image into the cover image to confirm authenticity and integrity of both image pixel data and image header. The scheme embeds the data into smooth blocks inside the ROI to achieve a watermarked image with low distortion. At extraction, the whole original image is fully recovered without the need for location map. The proposed method has been evaluated based on defined medical image watermarking requirements and compared to recent reversible watermarking approaches to verify its efficiency.

### 3 Proposed scheme

Conventional watermarking approaches based on DE embed one bit of watermark data into the difference value of two pixels. Locations of the pixels, used to encode the watermark, are required to detect/extract the watermark and reconstruct the reference image. The amount of additional information locating the relevant pixels reduces hiding capacity and increases potential for image distortion. Our watermarking approach encodes the watermark into smooth regions inside the ROI without needing a location map. This achieves high capacity watermarking with low distortion. The proposed method comprises three main steps; watermark creation, embedding, and extraction/verification. The parameters used in the embedding and extraction processes are listed in Table 1.

#### 3.1 Watermark creation

Several approaches can be used to generate watermark data for confirming the authenticity and integrity of medical images [13]. Some of authentication data is modified when the image is exchanged. In these cases, the embedded and recalculated authentication data are different,

**Table 1** The parameters used in the embedding and extraction processes

Parameter	Description
ROI	The ROI part of the input image
len	The length of the watermark data
fin_th	The final threshold used to identify the smooth blocks
OM	The original image
smb	The smooth blocks inside the ROI
wc	A binary array includes the encoded watermark
WM	The watermarked image
XM	The extracted image
wx	A binary array includes the extracted watermark

rendering authenticity confirmation impossible. This makes a careful selection of the authentication watermark is a necessity.

### 3.1.1 Authentication watermark

In addition to the image raw data, DICOM defines a structure for describing the image. This structure is located in the image's header and called metadata. DICOM metadata comprises tables of attributes which record key information including time of image acquisition, device parameters, imaging conditions, diagnosis result, and essential patient details such as the name, ID number, age, gender, weight, and height [26]. Some metadata fields are changed each time the image is distributed whilst others remain constant. Therefore, only information related to the patient and image (i.e. the constant data) must be used to ensure the authenticity. In our research, only essential metadata fields, which contain the patient information and data describing the image that do not change during distribution, were employed in the authentication watermark (*AW*) (Table 2). There is no necessity to utilize all columns, and only the value field is needed to create the watermark for ensuring the authentication.

### 3.1.2 Integrity watermark

The Digital Signature (*DS*) of the original medical image is calculated utilizing the Message Digest (*MD5*) algorithm. The *MD5* is a cryptographic hash function that generates a 128-bit Message Authentication Code (*MAC*). Any change to the image, either intentional or accidental, leads to change in the hash code. Comparing the base and retrieved codes enables identification of image manipulation [1]. In our research, the *DS* of the entire image is computed and encoded into the medical image to offer strict integrity watermark (*IW*).

### 3.1.3 Watermark compression

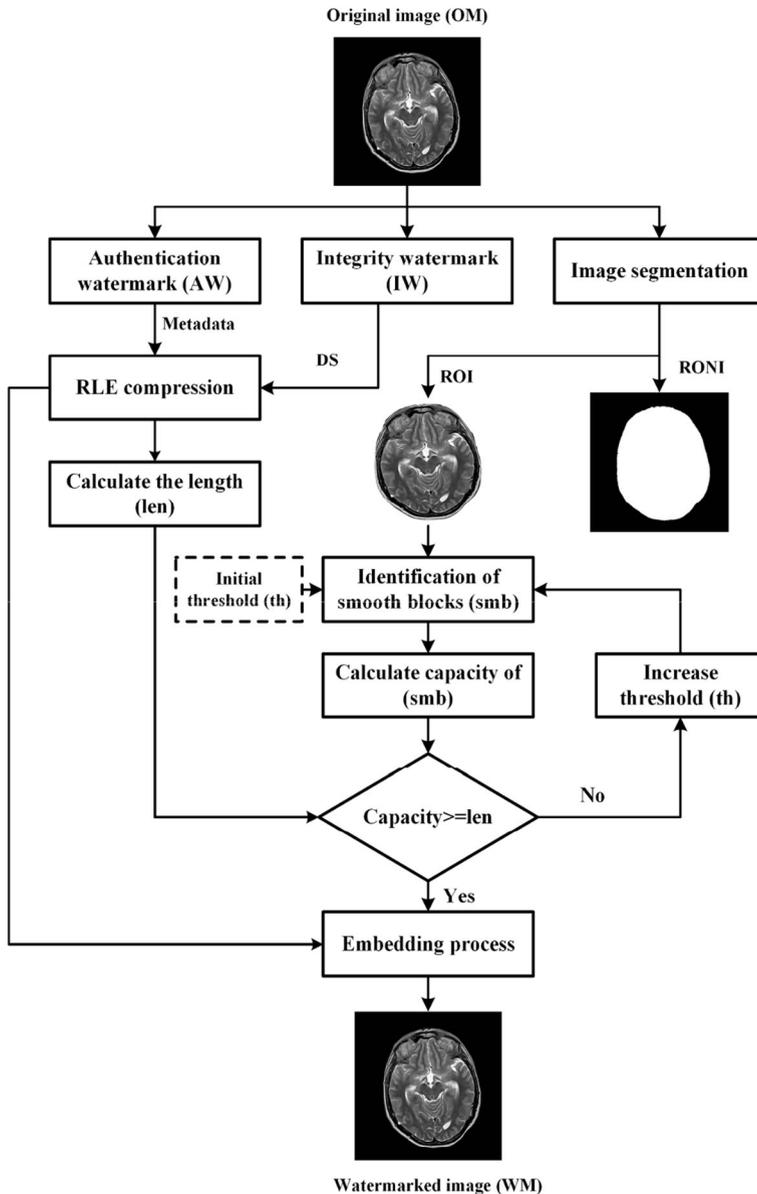
The two constructed watermarks (*AW* and *IW*) are concatenated and converted to binary form. To enhance the embedding capacity and reduce the distortion level, the watermark is compressed using Run Length Encoding (*RLE*). *RLE* is easy and quick to implement, making it a good alternative to other complex compression algorithms [30].

**Table 2** A section of metadata selected from a DICOM data dictionary [37] (These data do not relate to a real patient)

Tag	Description	VR	Value
0008,0020	Study Date	DA	01012018
0008,0030	Study Time	TM	103,045
0008,0060	Modality	CS	MR
0008,0070	Manufacturer	LO	SIEMENS
0008,0080	Institution Name	LO	Venice Hospital
0008,0090	Physician's Name	PN	Doctor Bellario
0010,0010	Patient Name	PN	Launelot Gobbo
0010,0020	Patient ID	LO	999,999
0010,0030	Patient Birth Date	DA	25,121,950
0010,0040	Patient Sex	CS	M
0018,0015	Body Part Examined	CS	Brain

### 3.2 Embedding process

The embedding process initially segments the cover image into ROI and RONI (Fig. 1). In this research, we considered the entire brain region as the ROI due to its importance in diagnosis. The smooth blocks inside ROI section are determined and the generated watermark is encoded into these blocks using a reversible watermarking method based on DE.



**Fig. 1** Process diagram for the embedding process. Starts by segmenting the cover image into ROI and RONI. Smooth blocks inside ROI are then identified. Watermark is encoded into the smooth blocks inside the ROI

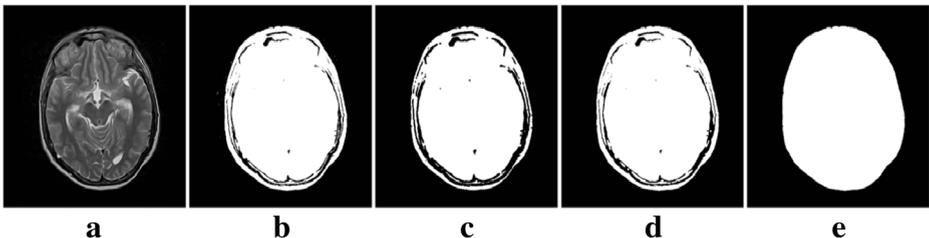
### 3.2.1 Image segmentation

Prior knowledge indicates that background intensity values of the brain MR slices are usually small compared to the intensity values in the foreground [18]. In our research, histogram thresholding was adopted as a segmentation technique to isolate background and identify the image ROI. This method is based on the thresholding values ( $T$ ). If the intensity value of a pixel is greater than  $T$  then the pixel is considered as a brain region (ROI), otherwise, it is assumed to be part of the background. The  $T$  value can be identified either manually or automatically by applying established approaches [19]. The  $T$  value was chosen experimentally (75) after applying a range of threshold values on many various images and visually evaluating these. A set of morphological operators, erosion, dilation and holes filling, are utilized to eliminate holes occurring in the segmented region (Fig. 2). Erosion is an operation used to decrease the size of the foreground objects and increase the size of the background. Dilation is an operation employed to increase the size of the foreground objects in binary images. A hole filling operator was applied to automatically fill the holes that were considered as background region in the binary image and surrounded by linked borders of foreground regions [43].

### 3.2.2 Smooth regions identification

Most medical images have a large smooth area, which is defined the regions that have little significant difference between the adjacent pixels intensity values, compared to the other images. Embedding the watermark into these regions is less noticeable to the human eye [6]. Consequently, the watermark was encoded in smooth areas inside ROI to decrease the degradation of the watermarked image. If adopting one of the existing techniques to determine the smooth regions, then when trying to identify the smooth regions to extract the encoded data, some of the identified smooth blocks will not precisely match the original blocks. Consequently, there is no guarantee that all pixels employed to discover the watermark will be similar to those utilized in the hiding process. This leads to the inability of the algorithm to extract the encoded data and recover the original image precisely.

We propose a simple new algorithm (Algorithm1) to determine the smooth regions inside the ROI of the medical image, such that when applying this algorithm to the watermarked image, it generates the same smooth blocks used at both embedding and



**Fig. 2** An example of MR slice segmentation, **a** Original, **b** Segmented, **c** Eroded, **d** Dilated, and **e** Filled holes. A binary matrix of one and zero values represents the foreground (ROI) and the background (RONI) respectively

**Fig. 3** An example of a  $3 \times 3$  block of pixels inside ROI which is individually treated and categorized as either smooth or non-smooth block

<b>P<sub>1</sub></b> (i-1,j-1)	<b>P<sub>2</sub></b> (i-1,j)	<b>P<sub>3</sub></b> (i-1,j+1)
<b>P<sub>4</sub></b> (i,j-1)	<b>P<sub>5</sub></b> (i,j)	<b>P<sub>6</sub></b> (i,j+1)
<b>P<sub>7</sub></b> (i+1,j-1)	<b>P<sub>8</sub></b> (i+1,j)	<b>P<sub>9</sub></b> (i+1,j+1)

extraction. This enables a precise extraction of the embedded data in the watermarked image without the need for any additional information (e.g. location map).

The algorithm segments the ROI into non-overlapping blocks of  $3 \times 3$  pixels which are separately evaluated and classified as either smooth or non-smooth blocks (Fig. 3). The absolute difference values between the corner pixels ( $P_1, P_3, P_7, P_9$ ) are calculated, and the average of these differences is computed and compared to the threshold value. The threshold value is increased based on the length of the watermark, created previously, to identify smooth blocks inside the ROI.

In Algorithm1, *ROI* denotes the ROI part of the image, *len* indicates the length of the watermark, and *fin\_th* represents the final threshold used to identify the smooth blocks (Table 1).

### 3.2.3 Watermark data encoding

After the image has been segmented to ROI and RONI and the smooth blocks inside ROI have been identified, the generated watermark data can be encoded. We extend Alattar [3] scheme using DE of five pixels, instead of quad pixels, to hide four bits of the watermark data in each of the five pixels ( $P_2, P_4, P_5, P_6, P_8$ ) of the identified smooth blocks. This keeps the corner pixels unchanged, which is required to extract the embedded data and recover the reference unmodified image without the need for any auxiliary information in the form of location map. The final threshold value and the length of the watermark are embedded into the RONI section using 1-bit per 2-pixels reversible watermarking algorithm [31].

For each identified smooth block, the embedding algorithm deducts the value of the center pixel ( $P_5$ ) from pixels ( $P_2, P_4, P_6, P_8$ ). Four new values are generated by encoding 4-bits of the watermark data into the differences values which previously calculated using the LSB technique. Finally, the inverse DE transform is applied to the generated new values, which carries the watermark bits, to produce the watermarked pixels.

In the embedding algorithm (Algorithm2), *OM* indicates the original image, *ROI* is the ROI part of the original image, *smb* denotes the smooth blocks inside the ROI, *we* is a binary array includes the watermark data, *len* indicates the length of the watermark (*we*), and *WM* represents the watermarked image (Table 1).

**Algorithm1** Smooth Region Identification*Inputs:* ROI, len*Output:* fin\_th

---

```

Step1:   th:=1;           //Assign an initial value to temporary threshold
Step2:   s:=0;           //Initiate the number of smooth blocks
Step3:   FOR (each 3x3 block in ROI) DO
//       Calculate the absolute differences values between the corner pixels (P1,P3,P7,P9) (Steps 4-9)
Step4:   dif1:=abs(P1-P3);
Step5:   dif2:=abs(P1-P7);
Step6:   dif3:=abs(P1-P9);
Step7:   dif4:=abs(P3-P7);
Step8:   dif5:=abs(P3-P9);
Step9:   dif6:=abs(P7-P9);
Step10:  d=average(dif1,dif2,dif3,dif4,dif5,dif6);           //Average of the difference values
Step11:  IF (d<=th) THEN           //Check if the average ≤ the temporary threshold
Step12:      s:=s+1;           //Increase the total number of smooth blocks
Step13:  END IF
Step14:  END FOR
Step15:  s:=s*4;           //Calculate the capacity of the smooth blocks
Step16:  IF (s<len) THEN           //Check if the capacity is enough for encoding the watermark
Step17:      th:=th+1;           //Increase the temporary threshold
Step18:      GO TO (Step3);           //Repeat the process of smooth blocks identification
Step19:  ELSE
Step20:      fin_th:=th;           //Assign the temporary threshold to the final threshold
Step21:  END IF

```

---

**3.3 Extraction and verification process**

Extraction and verification segments the watermarked image into ROI and RONI (Fig. 4). The final threshold and length of the embedded watermark are extracted from the RONI to identify smooth blocks inside the ROI. Concealed data is extracted from the pixels that have been employed in the embedding process, and the original pixels values are reconstructed.

Matching to the encoding process, the extraction algorithm deducts the value of the center pixel ( $P_5$ ) from pixels ( $P_2, P_4, P_6, P_8$ ) for each smooth block. 4-bits of the watermark data are retrieved and four new values generated by extracting the LSB from the differences values. Finally, the inverse DE transform is applied to the generated new values to reproduce the original unmodified watermarked pixels.

**Algorithm 2** Encoding Process

Inputs: OM, ROI, smb, we, len

Output: WM

---

```

Step1:   i:=1;      //Initiate the watermark counter
Step2:   FOR (each smb block in ROI) DO
//       Calculate the forward DE transform for the 5 pixels (P2,P4,P5,P6,P8) (Steps 3-7)
Step3:   v0:=floor((a*p5+b*p2+c*p4+d*p6+e*p8)/(a+b+c+d+e)) // Floor function indicates "the
//                                               greatest integer less than or equal
//                                               to" and a,b,c,d,e are integers

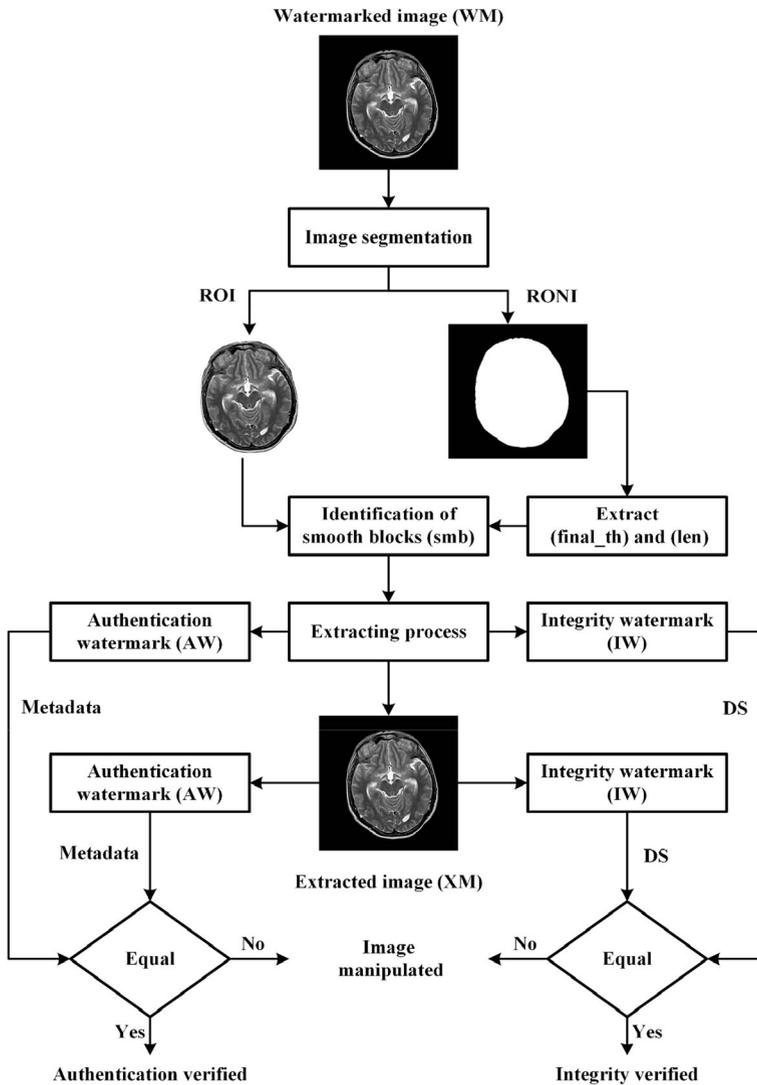
Step4:   v1:=P2-P5;
Step5:   v2:=P4-P5;
Step6:   v3:=P6-P5;
Step7:   v4:=P8-P5;
//       Embed 4 watermark bits into the differences values of the 5 pixels (Steps 8-12)
Step8:   nv0=v0;
Step9:   nv1=2*v1+we(i);
Step10:  nv2=2*v2+we(i+1);
Step11:  nv3=2*v3+we(i+2);
Step12:  nv4=2*v4+we(i+3);
//       Calculate the new pixels values (watermarked) (Steps 13-17)
Step13:  nP5=nv0-floor((b*nv1+c*nv2+d*nv3+e*nv4)/(a+b+c+d+e));
Step14:  nP2=nv1+P5;
Step15:  nP4=nv2+P5;
Step16:  nP6=nv3+P5;
Step17:  nP8=nv4+P5;
Step18:  i:=i+4; // Increase the watermark counter to encode the next 4 bits
Step19:  IF (i>len) THEN //Check the end of the watermark
Step20:      BREAK; //Exit
Step21:  END IF
Step22:  END FOR

```

---

In the extraction algorithm (Algorithm3), *WM* denotes the watermarked image, *ROI* indicates the ROI part of the watermarked image, *smb* identifies the smooth blocks inside the ROI, *fin\_th* is the final threshold, *len* indicates the length of the embedded watermark (*we*), *XM* represents the extracted image, and *wx* is a binary array includes the extracted watermark (Table 1).

The extracted watermark is decompressed using the same RLE decompression algorithm as for compression. It is divided into two watermarks; the authentication watermark (*AW*), and the integrity watermark (*IW*). These watermarks are compared to the recalculated metadata and DS of the extracted DICOM image to confirm authenticity and integrity of the image. This can be achieved by calculating the number of error and correct bits between the extracted and recalculated watermarks.



**Fig. 4** Process diagram for the extraction and verification process. Starts by segmenting the watermarked image into ROI and RON. Smooth blocks inside ROI are then identified. Watermark is extracted from the identified smooth blocks. The watermark is compared to the recalculated watermark of the extracted image to verify the authenticity and integrity of the image

### 4 Experimental results and discussion

To assess the performance of the proposed scheme, twenty-five brain MR scans in DICOM format (16bpp, 512 × 512 pixels) were used. Sixteen images are provided by the MRI unit of Al Kadhimiya Teaching Hospital (Iraq), from patients’ records for use in this research conducted at the University of Salford, UK (Fig. 5) [17]. Nine images are selected from a publically available and standardized medical images dataset downloaded from The Cancer Imaging Archive (TCIA) (Fig. 6) [12]. Several parameters have been used to conduct the experiment and evaluate the system

performance (Table 3). The Experimentation has been carried out using MATLAB R2016a working on MS Window 7 platform on a PC with Core i7–4790 Intel CPU and 16 GB RAM.

#### 4.1 Proposed system performance measurement

The proposed technique is assessed based on four principal requirements of image watermarking approaches: imperceptibility, reversibility, capacity, and robustness [32, 38]. Imperceptibility represents the highest requirement of watermarking systems. A digital watermark is called imperceptible if the original and watermarked images are perceptually indistinguishable. Imperceptibility is a factor of human cognition that needs to be appraised within the human context. We have conducted a visual assessment trial for 117 MR images in DICOM format [40]. These medical images have been modified by hiding a different amount of data to generate a range of images with various distortion levels. Five qualified radiographers evaluated the images through a relative Visual Grading Analysis (VGA) trial to determine the perceptual boundary, below which change is noticeable. This defines the level of modification that can be applied without perceptual distortion. The outcomes related to objective measures, includes PSNR for image fidelity. The results demonstrated that the modification of the images to a level of PSNR = 82 dB or better is unnoticeable to all observers, and modification level to a PSNR = 80 dB should not be noticeable in the vast majority of cases. Reversibility ensures the extraction of the watermark by precisely recovering the unmodified original image. The capacity refers to the number of watermark bits that can be concealed into the cover image. Robustness states the ability of resistance against different image processing operations such as rotating, resizing, adding noise, etc. Not all applications require robust watermark, in some applications, it is necessary to be fragile to detect alteration that can be applied to the images [38].

##### 4.1.1 Imperceptibility

Imperceptibility between the original, watermarked and extracted images has been measured utilizing the following commonly used metrics where  $MN$  is the images dimension, and  $I_o$  and  $I_w$  denote the original and watermarked images respectively.

**Peak Signal to Noise Ratio (PSNR)** It is a basic measure used to estimate the distortion amount between the original and watermarked images (Eq. 1). A higher PSNR value indicates lower distortion [38].

$$PSNR(I_o, I_w) = 10 \times \log_{10} \frac{MAX_I^2}{MSE} \quad (1)$$

Where  $MAX_I$  represents the highest possible pixel value of the input images and  $MSE$  is the Mean Squared Error between the original and watermarked images (Eq. 2).

$$MSE = \frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I_o(i, j) - I_w(i, j))^2 \quad (2)$$

**Structural Similarity Index (SSIM)** It is a Human Visual System (HVS) based measures to quantify the degradation in the structural information between two images. The SSIM approach compares the similarity of three factors: luminance, contrast, and structure (Eq. 3). It takes a value between  $-1$  and  $1$  where the value of  $1$  refers that the tested images are equal [32].

**Algorithm 3** Extraction Process*Inputs:* WM, ROI, smb, fin\_th, len*Outputs:* XM, wx

---

```

Step1:  i:=1;      //Initiate the watermark counter
Step2:  FOR (each smb block in ROI) DO
//      Calculate the forward DE transform for the 5 pixels (P2,P4,P5,P6,P8) (Steps 3-7)
Step3:      nv0:=P5;
Step4:      nv1:=P2-P5;
Step5:      nv2:=P4-P5;
Step6:      nv3:=P6-P5;
Step7:      nv4:=P8-P5;
//      Extract 4 watermark bits from the differences values of the 5 pixels (Steps 8-11)
Step8:      wx(i)=LSB(nv1);      //LSB is Least Significant Bit of binary representation of value
Step9:      wx(i+1)=LSB(nv2);
Step10:     wx(i+2)=LSB(nv3);
Step11:     wx(i+3)=LSB(nv4);
//      Calculate the new pixels values from the differences of the 5 pixels (Steps 12-16)
Step12:     v0=nv0+floor((b*nv1+c*nv2+d*nv3+e*nv4)/(a+b+c+d+e));      //a,b,c,d,e are the same
                                                                    integers used in the
                                                                    encoding process

Step13:     v1=floor(nv1/2);
Step14:     v2=floor(nv2/2);
Step15:     v3=floor(nv3/2);
Step16:     v4=floor(nv4/2);
//      Reconstruct the original pixels value (un-watermarked) from the new pixels values (Steps 17-21)
Step17:     P5= v0-floor((b*v1+c*v2+d*v3+e*v4)/(a+b+c+d+e));
Step18:     P2=v1+P5;
Step19:     P4=v2+P5;
Step20:     P6=v3+P5;
Step21:     P8=v4+P5;
Step22:     i:=i+4;      // Increase the watermark counter to extract the next 4 bits
Step23:     IF (i>len) THEN      //Check the end of the watermark
Step24:         BREAK;      //Exit
Step25:     END IF
Step26:     END FOR

```

---

$$SSIM(I_o, I_w) = \frac{(2\mu_{I_o}\mu_{I_w} + c1)(2cov + c2)}{(\mu_{I_o}^2 + \mu_{I_w}^2 + c1)(\sigma_{I_o}^2 + \sigma_{I_w}^2 + c2)} \quad (3)$$

$$\left\{ \begin{array}{ll} c1 = (k_1L)^2 & k_1 = 0.01 \\ c2 = (k_2L)^2 & k_2 = 0.03 \end{array} \right\}$$

Where  $\mu_{I_o}$  and  $\mu_{I_w}$  are the average of  $I_o$  and  $I_w$ , respectively,  $\sigma_{I_o}^2$  and  $\sigma_{I_w}^2$  are the variances of  $I_o$  and  $I_w$ , respectively. Cov is the covariance of  $I_o$  and  $I_w$ ,  $c1$  and  $c2$  are variables to stabilize the division with weak denominator, and  $L$  is the dynamic range of pixel values ( $L = 2^{\wedge}(\text{number of bits per pixels}) - 1$ ).

**Root Mean Squared Error (RMSE)** RMSE is the rooted value of the Mean Squared Error (MSE) and is mainly utilized to measure the reversibility of the watermarking technique (Eq. 4). RMSE value close to zero indicates low image distortion [39].

$$RMSE = \sqrt[2]{MSE} \quad (4)$$

**Image Fidelity (IF)** This metric measures the similarity between the original and watermarked images (Eq. 5). The value of IF equal to 1 indicates that the two images are identical [42].

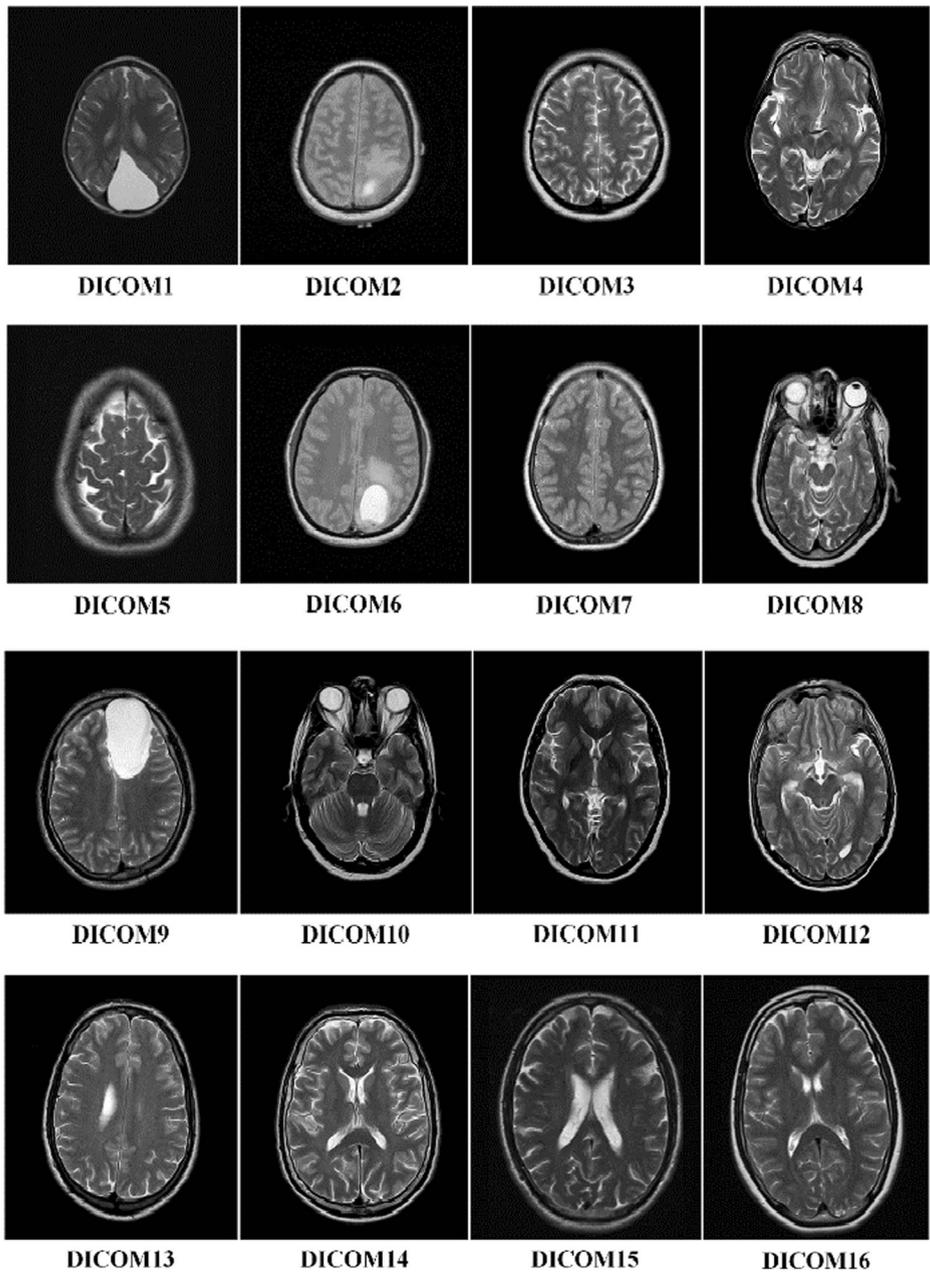
$$IF = 1 - \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I_o(i, j) - I_w(i, j))^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I_o(i, j))^2} \quad (5)$$

Examples of original DICOM images and their corresponding watermarked, extracted and the difference between the original and extracted images (Fig. 7) indicate that there is no opportunity to perceive any differences between the original and watermarked images. In addition, the results of imperceptibility between the various original, watermarked and extracted images, using PSNR, SSIM, RMSE and IF (Table 4) show that the PSNR values between the original and watermarked images are high, SSIM values are equal to one, RMSE values close to zero, and IF values are either one or very close to one. This indicates that the distortion of the watermarked image is very low and the watermark was encoded invisibly within the images. Therefore, the proposed method achieved the highest requirement of the digital watermarking schemes which is the imperceptibility.

#### 4.1.2 Reversibility

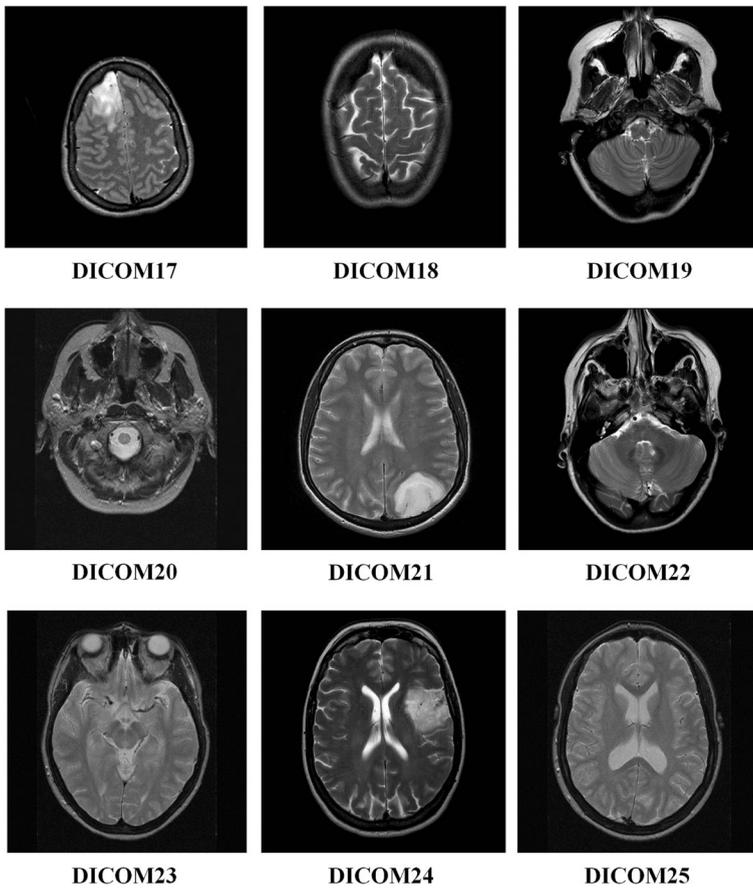
Reversibility of the proposed system has been assessed, in the extraction, for both the retrieved image and the extracted watermark.

**Image reversibility** The proposed technique need not require any additional information to detect/extract the encoded watermark and reconstruct the reference image. This is due to the ability of exactly identifying the same smooth blocks inside ROI in both the embedding and extraction process. The result demonstrates that there is no numerical difference between the original and extracted images (Fig. 7). PSNR values between the reference and extracted



**Fig. 5** The sixteen brain MR scans in DICOM format (16bpp,  $512 \times 512$  pixels) provided by the MRI unit of Al Kadhimiya Teaching Hospital (Iraq) [17] and used to assess the performance of the proposed method

images are equal to infinity, SSIM and IF values are equal to one, and RMSE values are equal to zero (Table 4). Consequently, after extracting the watermark, the extracted image is precisely identical to the reference image.

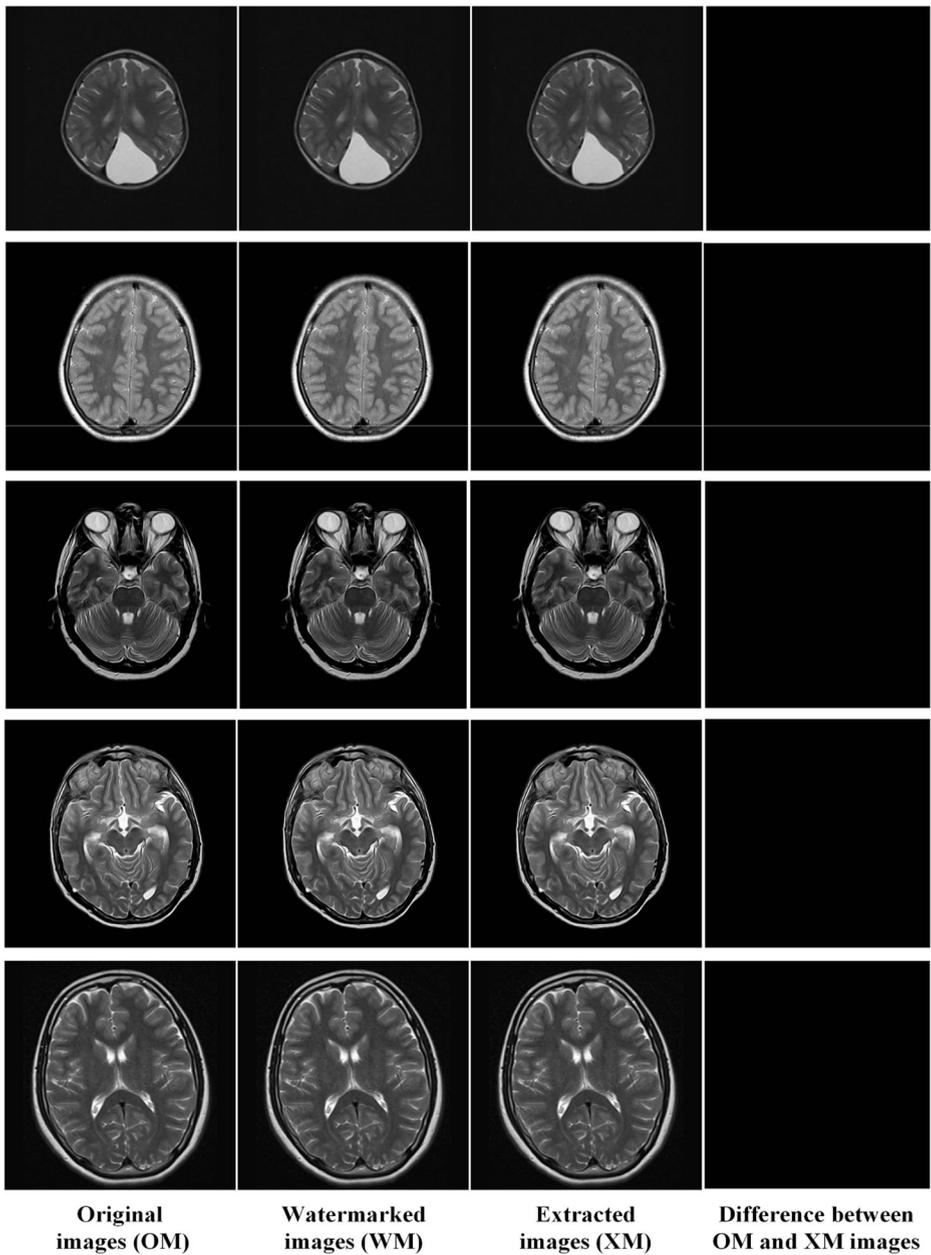


**Fig. 6** The nine brain MR scans in DICOM format (16bpp,  $512 \times 512$  pixels) selected from a publicly available and standardized medical images dataset downloaded from TCIA website [12] and used to assess the performance of the proposed method

**Watermark reversibility** The similarity between the embedded and extracted watermarks has been measured utilizing two commonly used metrics; Bit Error Rate (BER) (Eq. 6) and Accuracy Ratio (AR) (Eq. 7). These metrics calculate the number of error bits and the number of correct bits between the original and extracted watermark [42].

**Table 3** The parameters used to conduct the experiment and evaluate the proposed system performance

No	Parameter	Value
1	Images format	DICOM (16bpp)
2	Images modality	Brain MRI
3	Images size	$512 \times 512$ pixels
4	Watermark data	Authentication watermark (Metadata), Integrity Watermark (DS)
5	Performance evaluation criteria	Imperceptibility, reversibility, capacity, robustness
6	Attack/Manipulation type	Accidental, malicious



**Fig. 7** Examples of the original DICOM images and their corresponding watermarked, extracted and the difference between the original and extracted images. There is no visible difference between the original and watermarked images and no numerical difference between the original and extracted images

**Table 4** Imperceptibility between the original, watermarked, and extracted images using PSNR, SSIM, RMSE and IF metrics. These results indicate that the distortion of the watermarked image is very low and the watermark was encoded invisibly within the images

Images	Watermark length (bits)	Watermarked images				Extracted images			
		PSNR	SSIM	RMSE	IF	PSNR	SSIM	RMSE	IF
DICOM1	8224	99.94	1	0.0109	0.9999	∞	1	0	1
DICOM2	7496	94.93	1	0.0345	0.9998	∞	1	0	1
DICOM3	7328	94.57	1	0.0375	0.9998	∞	1	0	1
DICOM4	7288	96.72	1	0.0229	0.9999	∞	1	0	1
DICOM5	8296	94.10	1	0.0418	0.9999	∞	1	0	1
DICOM6	7440	97.78	1	0.0179	0.9999	∞	1	0	1
DICOM7	7472	94.91	1	0.0347	0.9998	∞	1	0	1
DICOM8	7368	93.23	1	0.0510	0.9998	∞	1	0	1
DICOM9	7352	96.15	1	0.0261	0.9999	∞	1	0	1
DICOM10	7384	95.69	1	0.0290	0.9999	∞	1	0	1
DICOM11	7384	97.77	1	0.0180	0.9999	∞	1	0	1
DICOM12	7576	96.39	1	0.0247	0.9999	∞	1	0	1
DICOM13	7352	98.51	1	0.0151	0.9999	∞	1	0	1
DICOM14	7312	97.29	1	0.0200	0.9999	∞	1	0	1
DICOM15	8256	97.68	1	0.0183	1	∞	1	0	1
DICOM16	8296	96.95	1	0.0217	1	∞	1	0	1
DICOM17	9336	95.12	1	0.0191	0.9998	∞	1	0	1
DICOM18	7648	95.17	1	0.0170	0.9999	∞	1	0	1
DICOM19	7688	92.50	1	0.0365	0.9998	∞	1	0	1
DICOM20	7424	93.07	1	0.0281	0.9998	∞	1	0	1
DICOM21	8104	92.18	1	0.0670	0.9998	∞	1	0	1
DICOM22	7632	95.22	1	0.0130	0.9999	∞	1	0	1
DICOM23	7432	94.04	1	0.0179	0.9999	∞	1	0	1
DICOM24	7624	96.73	1	0.0152	0.9999	∞	1	0	1
DICOM25	7456	94.64	1	0.0136	0.9999	∞	1	0	1

$$BER = \frac{EB}{TB} \quad (6)$$

$$AR = \frac{RB}{TB} \quad (7)$$

Where  $EB$  is the number of error bits,  $RB$  is the number of right bits and  $TB$  is the total number of watermark bits.

The results demonstrate that the BER values are equal to zero, and the AR values are equal to one. This indicates that the embedded watermark can be recovered at the extraction without any loss.

### 4.1.3 Capacity

Capacity of the watermarking system is determined by calculating the number of image pixels required for embedding the data (Eq. 8) [42].

$$Capacity = \frac{NP}{TP} \quad (8)$$

Where  $NP$  is the number of pixels required for embedding the watermark and  $TP$  is the total number of pixels.

Our proposed scheme encodes the watermark into the ROI part of the medical image. Therefore, the hiding capacity depends on the ROI size. The size of the watermark, used in this research for ensuring authenticity and integrity of the medical images, is approximately 1 KB. This indicates that the proposed scheme can encode the watermark even the size of ROI is 8% of the image size. The capacity of the proposed system is estimated by calculating the number of pixels required for embedding different magnitudes of data. Distortion level of embedding various payload into the sixteen DICOM images is measured using PSNR (Table 5). It is clear that the hiding capacity rises with increasing ROI size.

#### 4.1.4 Robustness

This research ensures the authenticity and integrity of DICOM images. Authenticity and integrity of the pixel data and header information of the watermarked image are confirmed, if and only if, the embedded watermark and original image can be retrieved correctly and exactly matched. Manipulations of the image data also corrupts the embedded watermark resulting in a mismatch between original and retrieved watermarks. PSNR and BER are used to assess reversibility and ability to recover the embedded watermark after applying image processing operations simulating both intentional and unintentional modifications. Malicious manipulations have also been applied including adding a new part to the image and removing an existing part from the image (e.g. lesion). The resultant PSNR values between the original and the extracted images are not equal to infinity, and BER values between the embedded and

**Table 5** PSNR values of hiding various payload (n/a refers to no space is available). It is clear that the hiding capacity and distortion level rise with increasing ROI size

Images	ROI size	Capacity 0.05bpp	Capacity 0.1bpp	Capacity 0.15bpp	Capacity 0.2bpp	Capacity 0.25bpp	Capacity 0.3bpp
DICOM1	27%	95.26	82.40	n/a	n/a	n/a	n/a
DICOM2	30%	90.10	80.90	n/a	n/a	n/a	n/a
DICOM3	35%	90.92	84.66	77.09	n/a	n/a	n/a
DICOM4	35%	93.05	86.01	74.57	n/a	n/a	n/a
DICOM5	36%	90.52	83.33	75.48	n/a	n/a	n/a
DICOM6	36%	93.25	85.68	74.23	n/a	n/a	n/a
DICOM7	37%	91.74	86.46	79.62	n/a	n/a	n/a
DICOM8	39%	89.84	84.10	78.31	n/a	n/a	n/a
DICOM9	41%	92.99	87.95	82.21	n/a	n/a	n/a
DICOM10	42%	91.76	85.48	79.26	n/a	n/a	n/a
DICOM11	42%	94.20	88.34	81.25	n/a	n/a	n/a
DICOM12	43%	93.05	87.08	80.85	n/a	n/a	n/a
DICOM13	46%	95.46	90.61	85.38	77.06	n/a	n/a
DICOM14	50%	94.00	88.35	82.21	75.34	n/a	n/a
DICOM15	59%	94.92	88.81	84.11	78.89	74.07	n/a
DICOM16	59%	93.71	87.56	82.56	76.54	70.63	n/a
DICOM17	26%	91.72	82.47	n/a	n/a	n/a	n/a
DICOM18	34%	91.08	84.88	n/a	n/a	n/a	n/a
DICOM19	40%	88.41	81.57	74.78	n/a	n/a	n/a
DICOM20	47%	89.29	84.20	79.95	n/a	n/a	n/a
DICOM21	49%	87.89	82.29	76.95	71.12	n/a	n/a
DICOM22	49%	91.12	85.04	79.57	72.57	n/a	n/a
DICOM23	50%	90.43	85.36	81.44	76.40	n/a	n/a
DICOM24	50%	92.99	87.59	82.72	75.65	n/a	n/a
DICOM25	51%	90.92	86.21	82.27	77.22	n/a	n/a

extracted watermarks are not equal to zero (Table 6). This demonstrates that the proposed system is fragile against various manipulations.

## 4.2 Comparison with existing approaches

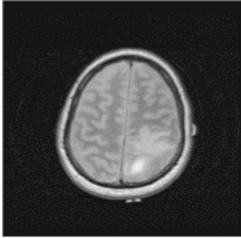
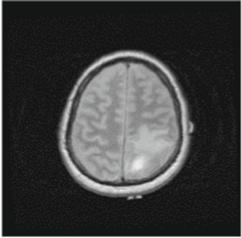
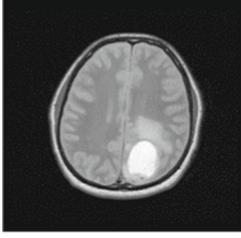
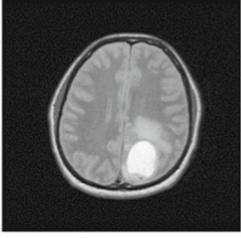
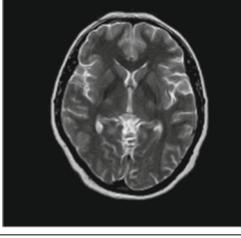
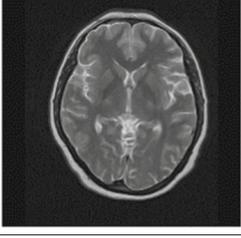
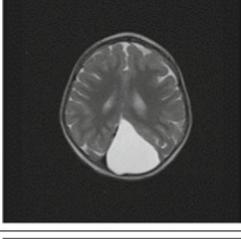
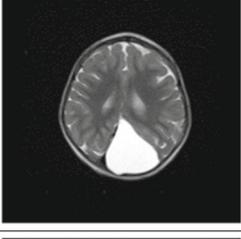
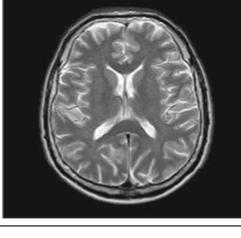
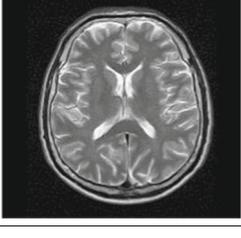
Performance of our proposed technique is compared with the existing reversible watermarking approaches. In general, reversible watermarking methods are implemented either in spatial or transform domains. Original pixel values of the whole image are utilized in the spatial domain. Therefore, hiding capacity is high, but robustness is low. In the transform domain, original pixels are changed and the whole image is not employed for encoding the data. Therefore, hiding capacity is lower but more robust.

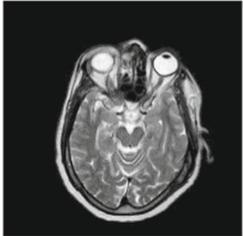
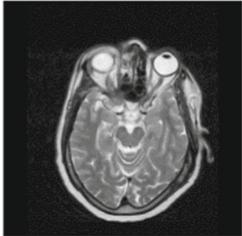
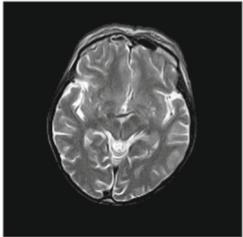
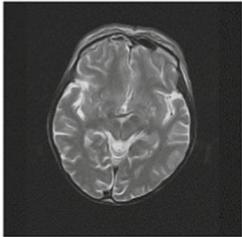
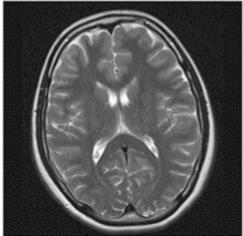
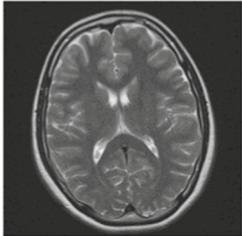
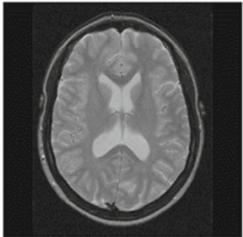
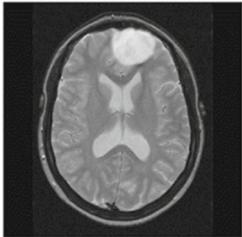
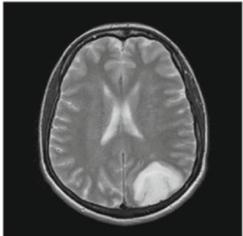
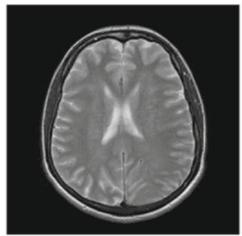
Many reversible watermarking approaches have been presented in literature to serve different applications. The performance of the proposed scheme is compared with reversible approaches applied to medical images to confirm authenticity and integrity (Table 7). In [14, 15], reversible watermarking methods are proposed in spatial domain. Embedding capacity is high, but several flaws can be observed; the ROI part of the image requires manual identification by a physician/clinician, and the ROI can only be recovered after removing the hidden data. Additional data is required for the extraction process. Transform based watermarking methods are proposed [9, 41, 42], and can recover the complete original image after extracting the watermark. However, image distortion is high in comparison to low embedding capacity. Moreover, approach [41] requires a location map for lossless recovery of the original image. In [16, 36], spatial domain based watermarking methods are presented. Watermark data is encoded into the ISB [36] and LSB [16] to ensure authenticity and integrity of the image. Although these schemes achieve high capacity, the distortion of watermarked images are high. Moreover, approach [16] needs additional information for extraction and the ROI can only be retrieved after extracting the watermark. Schemes [34, 46] utilize Histogram Shifting (HS) to hide the data. The cover image is automatically segmented into ROI and RONI and the original image is retrieved without any loss. However, the distortion of watermarked image is high in comparison to the amount of concealed data. Additionally, the location of pixels used to encode the data is required at the extraction. In [8] approach, the original unmodified image is recovered without the need for any additional information. A watermarked image with high visual quality is achieved, however, the scheme lacks evaluation of the embedding capacity. Moreover, a small number of images are used to assess the performance of the proposed watermarking technique.

Our proposed method is based on DE watermarking to ensure authenticity and integrity of both image pixel data and image header. The original medical image is segmented automatically into ROI and RONI. DICOM metadata and DS for the whole image is encoded into smooth region inside the ROI to protect the informative region of the image and make the distortion less visually perceptible. We evaluate image distortion through a clinical trial based on relative VGA to identify the perceptual distortion boundary. The obtained results demonstrate that the proposed scheme surpassed the other techniques in terms of visual image quality (Table 4). The complete original image is recovered at the extraction without the need for any auxiliary information. This contributed to controlling the capacity of the scheme and minimizing the distortion.

Furthermore, a DICOM15 image, which achieved highest embedding capacity, is used to compare the performance of our approach with other DE-based reversible watermarking schemes. These schemes comprise Tian [45], Alattar [3], Chiang, et al. [11], Al-Qershi and Khoo [6] (Scheme 1) and Al-Qershi and Khoo [6] (Scheme 2). DE-based watermarking

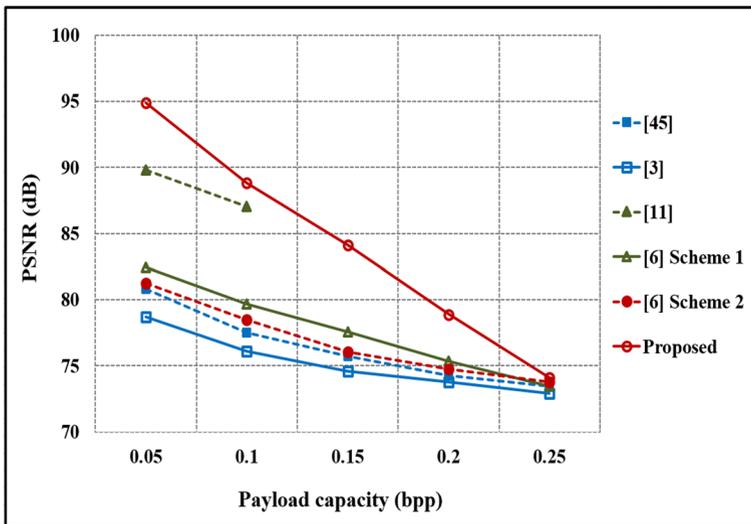
**Table 6** Reversibility evaluation of both original image and embedded watermark after applying various accidental and malicious manipulations. These results demonstrates that the proposed scheme is fragile against various manipulations.

Manipulation	Watermarked image	Manipulated image	PSNR	BER
Rotate ( $2^0$ ) & cropping			58.53	0.3837
Resize (90%)			Error	0.3827
Adjust brightness ( $\pm 5$ )			82.35	0.2451
Adjust intensity			7.06	0.3926
Gaussian filter ( $\sigma=0.5$ )			83.15	0.4981

Manipulation	Watermarked image	Manipulated image	PSNR	BER
Median filter			79.22	0.3857
Gaussian noise (SNR=5dB)			79.66	0.4967
Salt and pepper noise (d=0.0005)			76.45	0.4315
Adding region/lesion			17.70	0.4239
Removing region/lesion			16.92	0.4833

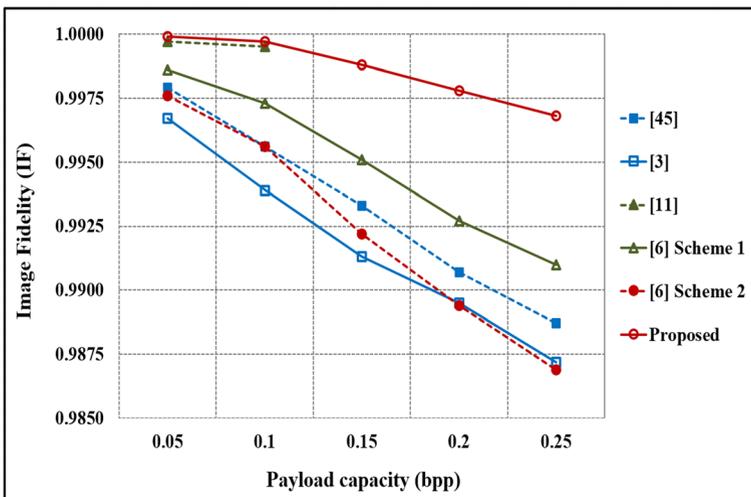
**Table 7** Performance comparison of the proposed scheme against approaches identified in literature. The results demonstrate that the proposed scheme surpassed the other techniques in terms of the visual image quality

Scheme	Year	Embedding technique	Embedding region	ROI segmentation?	Location map	Reversible?	Capacity	Visual quality
Das and Kundu [14]	2013	LSB	RONI	Manual	Yes	Only ROI	RONI dependent	Avg. PSNR = 43.85 dB Avg. MSSIM = 0.9904
Eswaraiah and Reddy [15]	2014	LSB	RONI Border pixels	Manual	Yes	Only ROI	Not presented	Avg. PSNR = 51.52 dB Avg. MSSIM = 0.9327
Roček, et al. [41]	2016	DT-CWT LSB	RONI	Automatic	Yes	Yes	RONI dependent	Avg. PSNR = 81 dB Avg. SSIM = 0.99997
Selvam, et al. [42]	2017	IWT DGT	Whole image	No	No	Yes	0.0625–0.255bpp	PSNR(60.42–65.23)dB SSIM(0.9832–1) RMSE(0.0321–0.0891) IF(0.9178–0.9836)
Parath, et al. [36]	2017	PTB conversion ISBS	Whole image	No	No	Yes	0.4–0.75bpp	PSNR(46.36–49.1)dB SSIM(0.9806–0.9899)
Gao, et al. [16]	2017	HS LSB	ROI RONI	Automatic	Yes	Only ROI	0.076–0.353bpp	PSNR(24.5–30.45)dB SSIM(0.9159–0.9829)
Scheme	Year	Embedding technique	Embedding region	ROI segmentation?	Location map	Reversible?	Capacity	Visual quality
Balasamy and Ramakrishnan [9]	2018	DWT	Whole image	No	No	Yes	0.1467bpp	PSNR = 49.01 dB SSIM = 0.9999
Yang, et al. [46]	2018	HS	ROI RONI	Automatic	Yes	Yes	0.01–3bpp	PSNR(14.74–23.06)dB SSIM(0.3504–0.9047)
Pan, et al. [34]	2018	HS	ROI RONI	Automatic	Yes	Yes	0.007–0.027bpp	PSNR(67.57–76.5)dB SSIM = 1
Atta-ur-Rahman, et al. [8]	2018	Residue with chaotically selected pixels	Whole image	No	No	Yes	Not presented	PSNR = 72.28 dB
Proposed scheme	–	DE	Smooth regions inside ROI	Automatic	No	Yes	Up to 0.255bpp	PSNR(70.63–99.94)dB SSIM = 1 RMSE(0.0109–0.0510) IF = 1



**Fig. 8** PSNR versus payload capacity for the proposed scheme, Tian [45], Alattar [3], Chiang, et al. [11], Al-Qershi and Khoo [6] (Scheme 1) and Al-Qershi and Khoo [6] (Scheme 2). The distortion is evaluated after hiding various payload magnitudes. The results indicate that the proposed algorithm achieves a watermarked image with a lower distortion

approaches, found in literature, can perform equally but most of them lack our simplicity [20, 25, 28]. Visual image quality is tested after hiding various payload magnitudes (0.05bpp, 0.1bpp, 0.15bpp, 0.2bpp, 0.25bpp). The results clearly signify that our proposed algorithm achieves a watermarked image with lower distortion in terms of PSNR (Fig. 8) and IF (Fig. 9) in comparison to other approaches.



**Fig. 9** IF versus payload capacity for the proposed scheme, Tian [45], Alattar [3], Chiang, et al. [11], Al-Qershi and Khoo [6] (Scheme 1) and Al-Qershi and Khoo [6] (Scheme 2). The distortion is evaluated after hiding various payload magnitudes. The results indicate that the proposed algorithm achieves a watermarked image with a lower distortion

## 5 Conclusion

In this paper, a novel blind reversible watermarking method is proposed for ensuring integrity and authenticity of brain MR images and detecting both accidental and malicious manipulations. The proposed scheme automatically segments the image into two sections; ROI and RONI. An extended reversible watermarking method based on DE technique is utilized to encode 4-bits of the watermark data in each smooth block of  $3 \times 3$  pixels selected from the ROI. The exact original images are retrieved after extracting the embedded watermark successfully. The need for a location map of the employed pixels is eliminated in the embedding and extraction processes. This enables, controls and facilitates maximizing hiding capacity whilst reducing image distortion. Based on the methodology of the DE technique, which hides the data in the difference values of pixels pairs, the proposed scheme encodes the watermark into the smooth blocks inside ROI. This makes the distortion less noticeable to the human visual perception and has been evaluated through visual trial based on the clinically recognized relative VGA technique to define a perceptual boundary, below which change is noticeable. This defines a clear metric to determine the level of modification that can be applied for encoding a known magnitude of payload data in an imperceptible manner.

Experimental results indicate that our proposed method yields superior performance to the other state-of-art schemes in terms of distortion level. It achieves excellent visual image quality regarding PSNR, SSIM, RMSE, and IF, and the proposed watermarking scheme is ideal for images with small ROI.

For future work, we recommend extending the proposed method to recover the manipulated regions, whilst operational trials are required before testing the watermark scheme in a fully operational PACS where the medical images are archived and retrieved.

**Acknowledgements** We would like to thank the ministry of higher education and scientific research (Iraq), for providing scholarship support to the first author of this research paper.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. Abd-Eldayem MM (2013) A proposed security technique based on watermarking and encryption for digital imaging and communications in medicine. *Egypt Inform J* 14:1–13
2. Alattar AM (2003) Reversible watermark using difference expansion of triplets. In: *International Conference on Image Processing (Cat. No.03CH37429)*, Barcelona, Spain, p 501–504
3. AM Alattar (2004) Reversible watermark using difference expansion of quads. In: *International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*, Montreal, Que., Canada, p 377–380

4. Alattar AM (2004) Reversible watermark using the difference expansion of a generalized integer transform. *IEEE Trans Image Process* 13:1147–1156
5. Ali AH, George LE, Zaidan A, Mokhtar MR (2018) High capacity, transparent and secure audio steganography model based on fractal coding and chaotic map in temporal domain. *Multimed Tools Appl* 77:31487–31516
6. Al-Qershi OM, Khoo BE (2011) High capacity data hiding schemes for medical images based on difference expansion. *J Syst Softw* 84:105–112
7. Arsalan M, Malik SA, Khan A (2012) Intelligent reversible watermarking in integer wavelet domain for medical images. *J Syst Softw* 85:883–894
8. Atta-ur-Rahman, Sultan K, Aldhafferi N, Alqahtani A, Mahmud M (2018) Reversible and fragile watermarking for medical images. *Comput Math Methods Med* 2018:1–7
9. Balasamy K, Ramakrishnan S (2018) An intelligent reversible watermarking system for authenticating medical images using wavelet and PSO. *Cluster Comput*:1–12
10. Celik MU, Sharma G, Tekalp AM, Saber E (2005) Lossless generalized-LSB data embedding. *IEEE Trans Image Process* 14:253–266
11. Chiang KH, Chang-Chien KC, Chang RF, Yen HY (2008) Tamper detection and restoring system for medical images using wavelet-based reversible data embedding. *J Digit Imaging* 21:77–90
12. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P et al (2013) The Cancer imaging archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 26:1045–1057
13. Coatrieux G, Lecornu L, Sankur B, Roux C (2006) A review of image watermarking applications in healthcare. In: *International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, USA, p 4691–4694
14. Das S, Kundu MK (2013) Effective management of medical information through ROI-lossless fragile image watermarking technique. *Comput Methods Prog Biomed* 111:662–675
15. Eswarajah R, Reddy ES (2014) Medical image watermarking technique for accurate tamper detection in ROI and exact recovery of ROI. *International Journal of Telemedicine and Applications* 2014:1–10
16. Gao G, Wan X, Yao S, Cui Z, Zhou C, Sun X (2017) Reversible data hiding with contrast enhancement and tamper localization for medical images. *Inf Sci* 385:250–265
17. Hasan AM, Meziane F (2016) Automated screening of MRI brain scanning using grey level statistics. *Comput Electr Eng* 53:276–291
18. Hasan AM, Meziane F, Kadhim MA (2016) Automated segmentation of tumours in MRI brain scans. In: *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC)*, Rome, Italy, p 55–62
19. Hasan AM, Meziane F, Aspin R, Jalab HA (2016) Segmentation of brain tumors in MRI images using three-dimensional active contour without edge. *Symmetry* 8:132
20. He W, Zhou K, Cai J, Wang L, Xiong G (2017) Reversible data hiding using multi-pass pixel value ordering and prediction-error expansion. *J Vis Commun Image Represent* 49:351–360
21. Khan A, Malik SA (2014) A high capacity reversible watermarking approach for authenticating images: exploiting down-sampling, histogram processing, and block selection. *Inf Sci* 256:162–183
22. Ko LT, Chen JE, Shieh YS, Scalia M, Sung TY (2012) "A novel fractional-discrete-cosine-transform-based reversible watermarking for healthcare information management systems. *Math Probl Eng* 2012:1–7
23. Ko LT, Chen JE, Shieh YS, Hsin HC, Sung TY (2012) Nested quantization index modulation for reversible watermarking and its application to healthcare information management systems. *Comput Math Methods Med* 2012:1–8
24. Kobayashi LOM, Furuie SS, Barreto PSLM (2009) Providing integrity and authenticity in DICOM images: a novel approach. *IEEE Trans Inf Technol Biomed* 13:582–589
25. Kumar V, Natarajan V (2016) Hybrid local prediction error-based difference expansion reversible watermarking for medical images. *Comput Electr Eng* 53:333–345
26. Larobina M, Murino L (2014) Medical image file formats. *J Digit Imaging* 27:200–206
27. Lee S-H, Lee E-J, Hwang W-J, Kwon K-R (2017) Reversible DNA data hiding using multiple difference expansions for DNA authentication and storage. *Multimed Tools Appl* 77:1–28
28. Lei B, Tan EL, Chen S, Ni D, Wang T, Lei H (2014) Reversible watermarking scheme for medical image based on differential evolution. *Expert Syst Appl* 41:3178–3188
29. Li F, Mao Q, Chang C-C (2018) Reversible data hiding scheme based on the Haar discrete wavelet transform and interleaving prediction method. *Multimed Tools Appl* 77:5149–5168
30. Liew SC, Liew SW, Zain JM (2013) Tamper localization and lossless recovery watermarking scheme with ROI segmentation and multilevel authentication. *J Digit Imaging* 26:316–325
31. Maity HK, Maity SP (2012) Intelligent modified difference expansion for reversible watermarking. *The International Journal of Multimedia & Its Applications (IJMA)* 4:83–95

32. Mousavi SM, Naghsh A, Abu-Bakar S (2014) Watermarking techniques used in medical images: a survey. *J Digit Imaging* 27:714–729
33. Nguyen T-S, Chang C-C, Huynh N-T (2015) A novel reversible data hiding scheme based on difference-histogram modification and optimal EMD algorithm. *J Vis Commun Image Represent* 33:389–397
34. Pan W, Bouslimi D, Karasad M, Cozic M, Coatrieux G (2018) Imperceptible reversible watermarking of radiographic images based on quantum noise masking. *Comput Methods Prog Biomed* 160:119–128
35. Parah SA, Sheikh JA, Ahad F, Loan NA, Bhat GM (2017) Information hiding in medical images: a robust medical image watermarking system for E-healthcare. *Multimed Tools Appl* 76:10599–10633
36. Parah SA, Ahad F, Sheikh JA, Bhat G (2017) Hiding clinical information in medical images: a new high capacity and reversible data hiding technique. *J Biomed Inform* 66:214–230
37. Piatnykh OS (2009) *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*. 2nd ed. Springer, Berlin
38. Qasim AF, Meziane F, Aspin R (2018) Digital watermarking: applicability for developing trust in medical imaging workflows state of the art review. *Comput Sci Rev* 27:45–60
39. Qasim AF, Meziane F, Aspin R (2018) A reversible and imperceptible watermarking scheme for MR images authentication. In: *Proceedings of the 24th International Conference on Automation and Computing (ICAC2018)*, Newcastle University, Newcastle upon Tyne, UK, 6–7 September 2018
40. Qasim AF, Aspin R, Meziane F, Hogg P (2019) Assessment of perceptual distortion boundary through applying reversible watermarking to brain MR images. *Signal Process Image Commun* 70:246–258
41. Roček A, Slavíček K, Dostál O, Javorník M (2016) A new approach to fully-reversible watermarking in medical imaging with breakthrough visibility parameters. *Biomed Signal Process Control* 29:44–52
42. Selvam P, Balachandran S, Iyer SP, Jayabal R (2017) Hybrid transform based reversible watermarking technique for medical images in telemedicine applications. *Optik* 145:655–671
43. Soille P (2013) *Morphological image analysis: principles and applications*. 2nd ed. Springer, Berlin
44. Tan CK, Ng JC, Xu X, Poh CL, Guan YL, Sheah K (2011) Security protection of DICOM medical images using dual-layer reversible watermarking with tamper detection capability. *J Digit Imaging* 24:528–540
45. Tian J (2003) Reversible data embedding using a difference expansion. *IEEE Trans Circuits Syst Video Technol* 13:890–896
46. Yang Y, Zhang W, Liang D, Yu N (2018) A ROI-based high capacity reversible data hiding scheme with contrast enhancement for medical images. *Multimed Tools Appl* 77:18043–18065



**Asaad F. Qasim** received the B.Sc. degree in Computer Science from Al-Rafidain University College, Iraq, in 2002 and the M.Sc. degree in Computer Science from the Iraqi Commission for Computers and Informatics, Informatics Institute for Postgraduate Studies, Iraq, in 2005. From 2006 to 2015, he was a Lecturer Assistance in the Ministry of Higher Education and Scientific Research in Iraq. He is currently pursuing the Ph.D. degree with the School of Computing, Science and Engineering, University of Salford, Greater Manchester, UK. His research interests include image processing, medical imaging, digital watermarking, integrity and authenticity in healthcare, and visual image quality assessment.



**Rob Aspin** is the Director of the School of Computer Science and Software Engineering, University of Salford. He is an experienced researcher and developer of virtual environments and their associated systems, with a background in both commercial development and academic research and development. His research interests include 3D computer graphics and virtual environments, advanced 3D information visualization, 2D and 3D medical imaging, modeling distribution for 3D graphics and virtual reality, 3D interaction, and non-real-time 3D rendering.



**Farid Meziane** is Professor in Data and Knowledge Engineering and the Director of the Informatics Research Centre in the School of Computing, Science and Engineering, University of Salford. He is the co-chair of the International Conference on Application of Natural Language to Information Systems and has served on the programme committees of over twenty conferences. He is on the editorial board of five international journals. His research interests include data and knowledge engineering, information extraction and retrieval, data mining and big data, natural language processing, and medical image processing.



**Peter Hogg** is Professor of Radiography at the University of Salford (UK) and Hanze University (Netherlands). He is also Research Dean in the School of Health Sciences, University of Salford. He qualified in diagnostic radiography in 1984, radiation protection in 1985 and nuclear medicine in 1986. Since 2009 he has been lead for the Diagnostic Imaging Research Programme within the University of Salford. His research interests include radiation dose optimisation in medical imaging and also cancer detection in mammography.

## Affiliations

Asaad F. Qasim<sup>1,2</sup> • Rob Aspin<sup>1</sup> • Farid Meziane<sup>1</sup> • Peter Hogg<sup>3</sup>

Rob Aspin  
R.Aspin@salford.ac.uk

Farid Meziane  
F.Meziane@salford.ac.uk

Peter Hogg  
P.Hogg@salford.ac.uk

<sup>1</sup> School of Computing, Science and Engineering, University of Salford, Salford, UK

<sup>2</sup> Ministry of Higher Education and Scientific Research, Baghdad, Iraq

<sup>3</sup> School of Health Sciences, University of Salford, Salford, UK