Egyptian Informatics Journal (2013) **14**, 1–13

Cairo University

Egyptian Informatics Journal

www.elsevier.com/locate/eij
www.sciencedirect.com

ORIGINAL ARTICLE

A proposed security technique based on watermarking and encryption for digital imaging and communications in medicine

Mohamed M. Abd-Eldayem *

Department of Information Technology, Faculty of Computers and Information, Cairo University, Egypt

Received 31 July 2012; revised 7 November 2012; accepted 21 November 2012
Available online 12 December 2012

KEYWORDS

Medical imaging;
Image authentication;
Image integrity;
Image compression;
Hash function;
Security

Abstract Nowadays; modern Hospital Data Management Systems (HDMSs) are applied in a computer network; in addition medicinal equipments produce medical images in a digital form. HDMS must store and exchange these images in a secured environment to provide image integrity and patient privacy. The reversible watermarking techniques can be used to provide the integrity and the privacy. In this paper, a security technique based on watermarking and encryption is proposed to be used for Digital Imaging and Communications in Medicine (DICOM). It provides patient authentication, information confidentiality and integrity based on reversible watermark. To achieve integrity service at the sender side; a hash value based on encrypted MD5 is determined from the image. And to satisfy the reversible feature; R–S-Vector is determined from the image and is compressed based on a Huffman compression algorithm. After that to provide confidentiality and authentication services: the compressed R–S-Vector, the hash value and patient ID are concatenated to form a watermark then this watermark is encrypted using AES encryption technique, finally the watermark is embedded inside the medical image. Experimental results prove that the proposed technique can provide patient authentication services, image integrity service and information confidentiality service with excellent efficiency. Concluded results for all tested DICOM medical images and natural images show the following: BER equals 0, both of SNR and PSNR are consistent and have large values, and MSE has low value; the average values of SNR, PSNR and MSE are 52 dB, 57 dB and 0.12 respectively. Therefore, watermarked images have high imperceptibility, invisibility and transparency. In addition, the watermark extracted from the image at the

* Corresponding author. Tel.: +20 966542974028.
E-mail address: mdayem@gmail.com.
Peer review under responsibility of Faculty of Computers and Information, Cairo University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.eij.2012.11.002

receiver side is identical to the watermark embedded into the image in the sender side; as a result, the proposed technique is totally reversible, and the embedded watermark does not affect the quality of the original image.

1. Introduction

Currently, most of Hospital Data Management Systems (HDMSs) and medical equipments are working in a computer network environment. Medical images are produced and stored in a digital form; moreover, they are exchanged through a computer network. These images are the most important entity in the healthcare diagnostic procedures because they are used to view features of patients such as anatomical cross-sections of internal organs and tissues, in addition they are used for physicians to evaluate the patient's diagnosis and monitor the effects of the treatment. Therefore, protecting medical images from an unauthorized use is an essential requirement. The most impotent security services required are patient privacy and medical image integrity, these security services can be provided using watermarking methods. A watermark is a part of information such as patient-ID and the image hash value that can be embedded in the image without corrupting this image.

The Digital Imaging and Communications in Medicine (DICOM) is the standard for formatting, storing and exchanging the medical images and associated information; moreover, DICOM support the connection of networked printers, such as laser imagers. Digital images could be acquired from diagnostic modalities such as: nuclear medicine, ultrasound, X-ray, CR, digital radiography, digitized film, video capture and hospital information system.

The watermarking techniques can be used to provide the integrity and the privacy; however, the diagnostic value in the medical image should not be changed. Lossless data hiding known as reversible watermarking embeds data within a digital image such that the original image can be completely restored. Therefore, many recent researches propose to use it for providing medical image integrity and patient privacy.

In this paper, a security technique based on watermarking and AES encryption methods is proposed to support DICOM security, the watermarking is reversible because the original image can be retrieved at the receiver side without any distortion. Experimental results prove that the proposed technique can provide patient authentication services, medical image integrity service and patient information confidentiality service with high efficiency.

The following sections of this paper are organized as follows: The related research works are summarized in Sections 2 and 3 describes the proposed medical image authentication technique based on the reversible watermarking method; the embedding and extraction processes are described in details. The DICOM standard and its file structure are described in Section 4; the experimental results are illustrated in Section 5, finally Section 6 summarizes the conclusion and future works.

2. Related research works

In this section, some of the recent medical image authentication techniques through watermarking are summarized. In [1] the image authentication and self-correction through an adaptive reversible watermarking technique are proposed. In this technique, the image is divided into two regions: Region Of Interest (ROI) and Region Of Non-Interest (RONI), it embeds the ROI into the RONI, and any modification of the image will be detected and could be self-restored back to the original image by extracting the ROI from the RONI. The ROI area is depending on the availability of clinical finding and its features in the medical image, and the RONI is the background or any area, where there is not any clinical finding. The pros of this technique are providing two levels of robustness by mixing a reversible watermarking method and a robust watermarking method. This watermarking method provides the initial level of robustness of the watermark extraction process against JPEG compression; a digital signature derived from the ROI, and an authenticity code is concatenated to form a primary code to be embedded inside the RONI using the robust watermarking method. The reversible watermarking technique provides the second level of robustness by embedding another code into the ROI; this code is determined for the whole image (RONI and ROI).

This proposed technique is used for a specific type of medical images that is Magnetic Resonance (MR) images; this type of medical images is very simple to identify RONI and ROI; therefore, this proposed technique is unable to authenticate other types of medical images that their RONI and ROI are hard to be separated.

In adaptive data hiding scheme for medical images using integer wavelet transform [2]; integer wavelet transform hiding technique is used for embedding the multiple watermarks by decomposes the cover image to obtain the wavelet coefficients. Before watermark embedding process; an adaptive threshold is determined for each block; it uses iterative optimization of threshold for compression and expansion process. It avoids histogram pre and post-processing; therefore, its pros are reducing the histogram processing overhead and keeping the distortion small between the watermarked and the original images. The cons of this technique are: low imperceptibility values at normal embedding capacity (bad tradeoff between robustness and capacity) and it is not applied to color images (it is applied only for grayscale images).

A multiple block based authentication watermarking for distribution of medical images is proposed in [3], it provides an active method of authentication for the efficient distribution of images, and this technique suggests a new method using fragmentation of the watermark information content of images. Medical imaging modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and the structure of tissues contain a large amount of clinical information. Therefore, it is important to provide authentication for the safety of fragmented blocks. The proposed technique is based on the secure encryption watermark, but removes the problem of independence block wise of existing methods. This technique suggests

to merge multiple signatures that were created through square blocks and blocks of fragmentation, two types of signatures shared to remove the block wise independence. The advantage of this proposed technique being able to detect the location of the modification; therefore, it neglects the tampers if the modifications are located in an RONI.

A medical image authentication based on lossless watermarking is proposed [4]; it is used for interleaving patient information and message authentication code with images using lossless compression. At embedding process the authentication code of the image using MD5 algorithm is calculated; the authentication code and patient information are concatenated then encrypted. Least Significant Bits (LSBs) of all pixels are selected and compressed using Run Length Encoding (RLE) lossless compression algorithm. The compressed string and the encrypted string are concatenated and inserted into the LSB locations by adding blanks if necessary. Before embedding process the patient information is encrypted; therefore, this technique provides a high level of security for the patient information. This proposed technique inherits the disadvantage of the LSB embedding process that is changing the statistical property of the cover image; therefore, the hiding process can be detected easily by computer systems.

In [5] a blind watermarking based on wavelet transform is proposed for medical image management, it hides the Electronic Patient Record (EPR) in the image: to protect patient information, to save storage space and to reduce transmission overheads. It embeds EPR data as a watermark in the Discrete Wavelet Packet Transform (DWPT) of the image. This proposed technique enhances the robustness by encoding EPR data using BCH error correcting code. The disadvantages of this technique are that it is purely implemented for grayscale images (not for color images), and it has been low embedded capacity. The embedded process hides only one bit per a block of pixels with size 4×4 pixels, and the error correcting code reduces the actual capacity to be less than one bit per 4×4 block of pixels.

In [6] a robust fragile watermarking technique is proposed to provide copyright protection and content authentication of medical images. It authenticates the CT scan images of the thorax area against distortions. It separates a ROI and RONI from the image. By isolating the actual lung parenchyma; this technique increases the embedding capacity of a CT scan image; it embeds a watermark only in RONI; therefore, it does not compromise the diagnostic value of the image. For embedding the watermark; it utilizes the spatial domain watermarking and LSB replacement method. The cons of this technique are it is devoted to a specific type of medical images; in addition, its robustness require to be improved.

A watermarking framework based on wavelet-domain is proposed [7], it proposes a robust reversible watermark embedding and extraction procedure through histogram shifting and clustering. It provides good performance in terms of reversibility, robustness and invisibility, but the embedded capacity is less than 4×10^{-3} bpp. It is applicable in practice to many types of medical images; however, it is tested using a limited number of grayscale images; therefore, it is required to be tested using enough number of grayscale and color images.

A blind image watermarking technique based on Contourlet (CN) transform is proposed for the medical data-management scheme [8], It is robust against high JPEG and JPEG2000 compression, and it can provide information security, content authentication, safe archiving and controlled access retrieval. In this proposal, an original image is decomposed based on CN transform, then the watermark is embedded inside the image using the low pass such that the embedded watermark can be extracted in a blind manner, finally the image is reconstructed based on the inverse of the CN transform to get the watermarked image. It can be used during a medical image acquisition process to provide authenticity, integrity and confidentiality, but the embedded capacity is very low it is less than 0.0053 bpp.

A reversible watermarking scheme based on image classification and histogram shifting is proposed [9]. In this scheme, each part of the image is watermarked with the most adapted lossless modulation between: Pixel Histogram Shifting (PHS) or Dynamical Error Histogram Shifting (DEHS); therefore, at first a reference image is created to identify the most efficient watermarked method: PHP or DEHS for each image part. The watermark embedding and extraction are implemented based on the reference image. DEHS dynamically shifts predicted-errors between the image and its reference image. This technique can embed high capacity with low distortion.

In [10] a reversible watermark based on Quantization Index Modulation (QIM) is proposed to be applied to healthcare information management systems, the QIM-based watermarking is used to reconstruct the identical original image; the capacity of the watermark is increased to be one-fourth of that of the cover image. Its architecture and algorithms are simple; it can be easily implemented. However, it is tested using only grayscale images; accordingly, it is required to be tested using color images.

In [11] a security technique based on encryption, and watermarking is proposed to protect medical images; it enables access to the outcomes of the encrypted image integrity and of its origins. With this technique, the RC4 stream ciphers and two substitute watermarking methods are combined; these two watermarking methods are the LSB and the QIM methods. In the embedding process, the watermarking and encryption are conducted jointly; therefore, in the extraction process, the watermark extraction and decryption can be applied independently. This technique can achieve a large embedded capacity in the spatial domain (0.5 bpp) with a high Peak Signal to Noise Ration (PSNR) that is greater than 49 dB. Due to using the LSB watermarking method; the statistical property of the watermarked images is changed; accordingly, the hidden information can be detected by the attacking computer system.

An adaptive dual blind watermarking scheme is proposed for medical images [12], it automatically selects the ROI and embeds the watermark with different embedding strength in ROI and RONI; it embeds watermark bits in singular values within the low-pass sub-band in the CN domain; therefore, it is more efficient and robust than embedding within the wavelet domain. This technique can be applied to DICOM image format; it has large PSNR and it satisfies high transparency for its watermarked images, but the invisibility could be enhanced. This technique is tested only using CT and MR images; therefore, it required to be tested using other types of images.

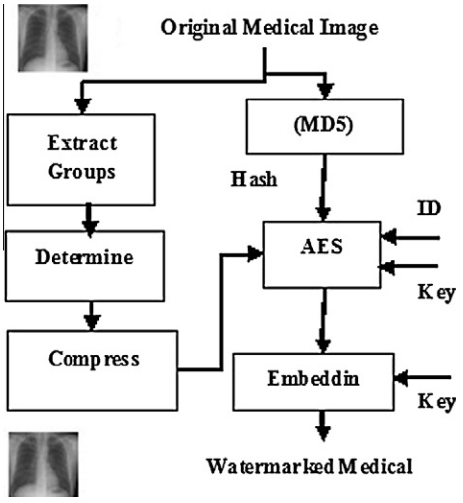

Figure 1 Embedding process.

3. The proposed technique for medical image integrity based on watermarking method

In this section, the details of the proposed technique are described, it proposes the medical image integrity for DICOM based on reversible watermarking and encryption methods. The watermark embedding process, watermark extracting process, R–S-Vector compression and the analysis of the technique performance are described in details.

3.1. Watermark embedding process

In embedding process, the algorithm achieves image integrity and authentication by adding the watermark to the image according to the following steps (Fig. 1):

(i) Extract groups.
(ii) Determine R–S-Vector.
(iii) Compress R–S-Vector.
(iv) Calculate the MD5 hash value of the image.
(v) Add the MD5 value to the compressed R–S-Vector and patient ID to get a watermark.
(vi) Encrypt the watermark using AES and Key 1.
(vii) The watermark is embedded by modifying the image using the watermark and key2.

These steps are explained in the following subsections (see Fig. 2).

3.1.1. Extract groups

In this step, the image is divided into groups; each group consists of four pixels, and it will be represented as a singular value. Discriminating and Flipping functions must be defined before identifying this single value of each group.

- Discriminating Function (f)

Discrimination function is used to describe the state of the group, and it is calculated based on

$$f(\text{Group}) = \sum |x_{i+1} - x_i| \quad (1)$$

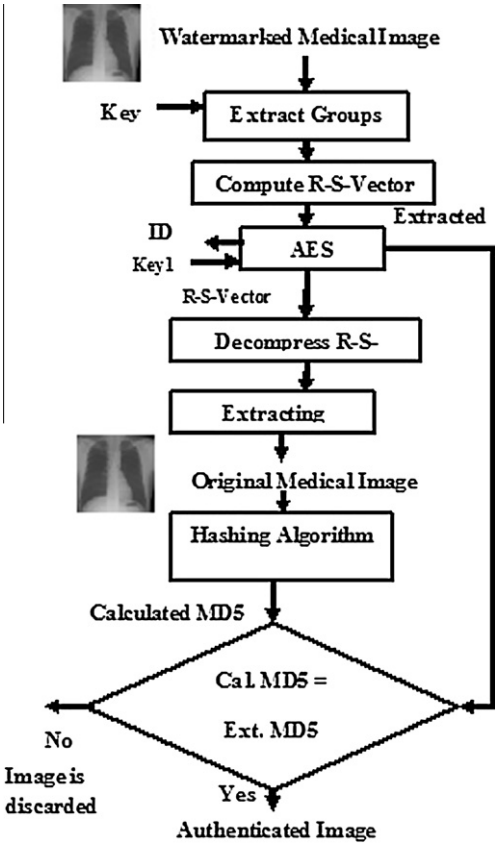

Figure 2 Extraction process.

where Group $= \{x_1, x_2, x_3, x_4\}$, x_i is the value of the pixel i in the current group.

- Flipping Function (F)

Flipping function is used to modify the pixel value by flipping the least significant bit. For example; if the value of a pixel before using this function equals 234 (that is represented in the binary system as 1110 1010). Then after using the function; the least significant bit will be flipped to be "1" instead of "0"; accordingly, the value of this pixel will be 235 (that is represented in the binary system as 1110 1011).

The discrimination function is calculated for each group before (f_{Before}) and after (f_{After}) using the flipping function; and the state of each group can be determined as follows [13–15]:

- **R** Group (**R**egular group): if $f_{\text{After}} > f_{\text{Before}}$
- **S** Group (**S**ingular group): if $f_{\text{After}} < f_{\text{Before}}$
- **U** Group (**U**nused group): if $f_{\text{After}} = f_{\text{Before}}$

For example; if the values of four pixels are (128, 129, 130, 128), that is represented in the binary system as (1000 0000, 1000 0001, 1000 0010, 1000 0000), the value of the discrimination function for this group of pixels is calculated according to Eq. (1):

$$F_{\text{Before}}(\text{Group}) = \sum(|129 - 128| + |130 - 129| + |128 - 130|)$$

$$F_{\text{Before}}(\text{Group}) = \sum(1 + 1 + 2) = 4$$

The values of the two middle pixels of the group will be modified using the flipping function. Accordingly, the values of the four pixels after using flipping function will be (128, 128, 131, 128) that is represented in the binary system as (1000 0000, 1000 0000, 1000 0011, 1000 0000). The value of the discrimination function for this group of pixels is determined using Eq. (1):

$$F_{After}(Group) = \sum(|128 - 128| + |131 - 128| + |128 - 131|)$$

$$F_{After}(Group) = \sum(0 + 3 + 3) = 6$$

The state of this group of pixels is **R** Group because of its f_{After} > its f_{Before}; its f_{After} = 6 while its f_{Before} = 4.

3.1.2. Creating R–S-Vector

Each group of pixels has a single value: **1** for **R** (Regular group), **0** for **S** (Singular group) and **−1** for **U** (Unused group). The unused groups will be ignored because they are not affected by the flipping function, therefore, the R–S-Vector consists of a stream of bits (zeros and ones), and each bit represents the state of a group of pixels in the image.

3.1.3. Compressing the R–S-Vector

The compression algorithm that is used to compress the R–S-Vector must satisfy the following criteria:

1. Lossless compression: to restore the original R–S-Vector.
2. Good compression ratio for this type of data.
3. Ability to compress a binary data (stream of bits).
4. Ability to compress a random data.

3.1.4. Calculating hash value of the image

The first objective of this research is to provide the medical image integrity service; a hash value of the image is determined using the MD5 hash function [16,17]. MD5 is proposed to be used because it produces Message Authentication Code (MAC) with size equal only 128 bits, and this is the least hash code size. Therefore, the saved size due to the compression of R–S-Vector will be sufficient to embed the MAC and the patient ID. Please note that the MAC will be encrypted to protect it against attacks.

3.1.5. Creating the watermark

The MD5 hash value, patient-ID and the compressed R–S-vector of the original image are concatenated, and then they are encrypted using AES encryption technique [16,18] to create the watermark; this watermark is embedded into the image. MD5 provides image integrity, patient-ID provides patient-authentication, and the AES is used to provide confidentiality for the patient information. AES is used to encrypt the watermark using a private key shared between the sender and the receiver (key1) to provide confidentiality.

Using AES that is a symmetric encryption technique does not provide sender authentication because the encryption key must be shared between the sender and the receiver. However, it provides the authentication for the users (for examples the doctors) of the medical system; It allows only the doctors and specialists to check the images and medical history of their patients, and this is used in modern medical systems.

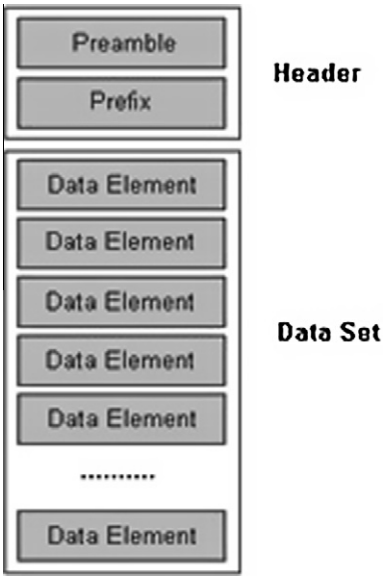

Figure 3 DICOM file format.

3.1.6. Modifying the original image according to the watermark

In this step, the watermark is embedded into the image; the state of each group of pixels of the original image is modified using the flipping function to represent one bit of the watermark; the output image of this step is called the watermarked image. It is optional to select the group of pixels randomly using key2 as a seed of the random number generator, this to distribute the watermark randomly inside the image.

3.2. Watermark extracting process

In Extraction process (Fig. 3), the original image is retrieved, and the watermark is extracted from the watermarked image; this process consists of the following steps:

1. Extract groups. (4 pixels per Group)
2. Determine The R–S-Vector and extract the encrypted watermark.
3. Decrypt the watermark using AES.
4. Extract the hash value (MD5), patient ID and the R–S-Vector of the original image.
5. Decompress the R–S-Vector.
6. Extract the original image (non-watermarked image).
7. Calculate the hash value (MD5) of the extracted original image.
8. Compare the calculated hash value (step5) and the extracted hashes value (step2), if they are equal, the image is authenticated, and it has right integrity, else the image is discarded because its integrity is broken.

Extract groups and Create R–S-Vector processes will be done as in the embedding process. The encrypted watermark consists of the compressed R–S-Vector and the hash value (MD5) of the original image; identify the compressed R–S-Vector and the hash value (MD5). The AES is used to decrypt the watermark using (key1).

Then the decrypted R–S-Vector is decompressed using Extended Huffman algorithm. The original image is extracted

by modifying the groups' status of the watermarked image to become identical to the decompressed R–S-Vector. The MD5 Hash value of extracted non-watermarked image is calculated; it is compared with the decrypted MD5 hash value, if they are equal, then the image is not modified, and it is validated, otherwise the two hash values are different, therefore, the image was modified, and it is no longer authenticated. If the image is authenticated, it is proven that its integrity is right and the encryption process is implemented with a legal user within the medical system because he has the private key, in addition the identity of the patient is identified using the extracted patient ID.

3.3. R–S-Vector compression process

It is required to compress the R–S-Vector by an efficient way to provide free space for embedding the watermark. The efficiency of the compression process depends on the nature of the data to be compressed and the bias among its symbols. More bias among the symbols of the data leads to a high compression ratio. Therefore, if the R–S-Vector has the adequate bias between the two symbols (ones for Regular groups, and zeros for Singular groups), then the compression process will provide sufficient space for embedding the watermark. The embedding capacity (Cap) can be calculated according to Eq. (2) [13,15]:

$$\text{Cap} = N_R + N_S - |C| \tag{2}$$

where (N_R) is the number of the Regular groups in the image, (N_S) is the number of the Singular groups in the image and ($|C|$) is the length of the compressed R–S-Vector. An ideal lossless context-free compression scheme (the entropy coder) would compress the R–S-Vector consisting of ($N_R + N_S$) bits according to Eq. (3) [15]. The target is to maximize embedding capacity by minimizing the compressed R–S-Vector length ($|C|$).

$$-N_R \log\left(\frac{N_R}{N_R+N_S}\right) - N_S \log\left(\frac{N_S}{N_R+N_S}\right) \text{bits} \tag{3}$$

From Eqs. (2) and (3), a theoretical estimate of the upper bound for the real capacity (**Cap**′) can be calculated according to [15]

$$Cap' = N_R + N_S + N_R \log\left(\frac{N_R}{N_R+N_S}\right) + N_S \times \log\left(\frac{N_S}{N_R+N_S}\right) \tag{4}$$

Practically, the flipping function was used to modify the LSB of the four pixels; the number of unused groups is increased. Therefore, in the proposed technique, it modifies only the two middle pixels of the group; this leads to increase the value of (NR + NS). In addition, the total number of groups is increased due to using joint groups. This means the two middle pixels of each group (each group has four pixels) are only belonging to a unique group; however, Least Significant Pixel (LSP) belongs to both of this group and the previous group, also Most Significant Pixel (MSP) belongs to both of this group and the next group. For example, if we have six pixels [p10, p11, p12, p13, p14, p15, p16, p17, p18, p19], then the first group consists of [p10, p11, p12, p13]. The second group includes [p13, p14, p15, p16], and the third group includes [p16, p17, p18, p19]. This method can embed one bit per three pixels; therefore, the embedded capacity is around 0.3 bpp.

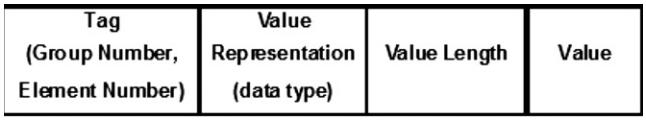

Figure 4 Data element structure.

To reduce ($|C|$), an efficient lossless compression algorithm is proposed to be used to compress R–S-Vector as described in the following paragraphs, in addition MD5 hash function is used because it creates the lows hash size. Therefore, the proposed technique provides adequate capacity.

The compression algorithm used to compress R–S-Vector must be lossless compression and has a suitable compression ratio to add the MD5 hash value and Patient-ID. In this research the ability of Run Length Coding, LZ77, LZ78, LZW, Huffman and Adaptive Huffman lossless compression algorithms [19–21] to compress the R–S-Vector is tested.

Run Length Coding has failed to satisfy a sufficient compression ratio because it depends on the slow change rates in the raw data symbols while the R–S-Vector data have a random fashion. In addition; the LZ77, LZ78 and LZW compression algorithms cannot provide an adequate compression ratio to compress S–R-Vector because they depend on a property that some patterns have great probability to appear in the raw data; these patterns can be used to build a codebook, and then this codebook is saved with the indices of the raw data itself. Adaptive Huffman is a statistically based compression algorithm depending on adapting property; for each symbol appears in the stream its probability ratio is increased, and its code length is decreased, if a symbol appears more and more its probability ratio becomes high, and its code length becomes short; these codes are assigned using a Codes Tree. Adaptive Huffman cannot compress S–R-Vector effectively. The Huffman algorithm depends on statistical calculations; distinctive symbols according to the probability of their appearance in the data stream can have a code with different length. A symbol with high appearance probability will have shorter code than other symbols with low appearance probability. The R–S-Vector consists of a stream of bits; consequently, to use a Huffman's algorithm, R–S-Vector is converted into a stream of symbols, symbols with size of 4 and 8 bits are used in the compression process. If the symbols with size 8 bits are used the R–S-Vector is divided into blocks with size of 8 bits, unfortunately this segmentation gave an

Table 1 Reults of Grayscale DICOM File.

	BER %	SNR (dB)	MSE	PSNR (dB)
GM1	0	49.96	0.148	56.43
GM2	0	51.56	0.121	57.30
GM3	0	64.17	0.106	57.88
GM4	0	48.86	0.112	57.64
GM5	0	71.95	0.113	57.60
GM6	0	44.45	0.1	58.13
GM7	0	72.92	0.156	56.20
GM8	0	49.13	0.136	56.80
Average	0	56.63	0.124	57.25
Standard deviation	0	11.29	0.020	0.70

insufficient compression ratio, because it produces a codebook with enormous size. When symbols with size 4 bits are used to compress R–S-Vector, fortunately it provides a sufficient compression ratio to compress the R–S-Vector, this is because the 4 bits block provides a small size codebook with only 16 symbols while 8 bits block provides a huge size codebook with 256 symbols. Therefore, it is proposed to use a Huffman compression algorithm with 4 bits symbol.

4. DICOM standard

The Digital Imaging and Communications in Medicine (DICOM) is the standard for formatting, storing and exchanging medical images and associated information. Digital images could be from diagnostic modalities such as Nuclear Medicine, Ultrasound, X-ray, CR, digital radiography, digitized film, video capture and hospital information system. DICOM

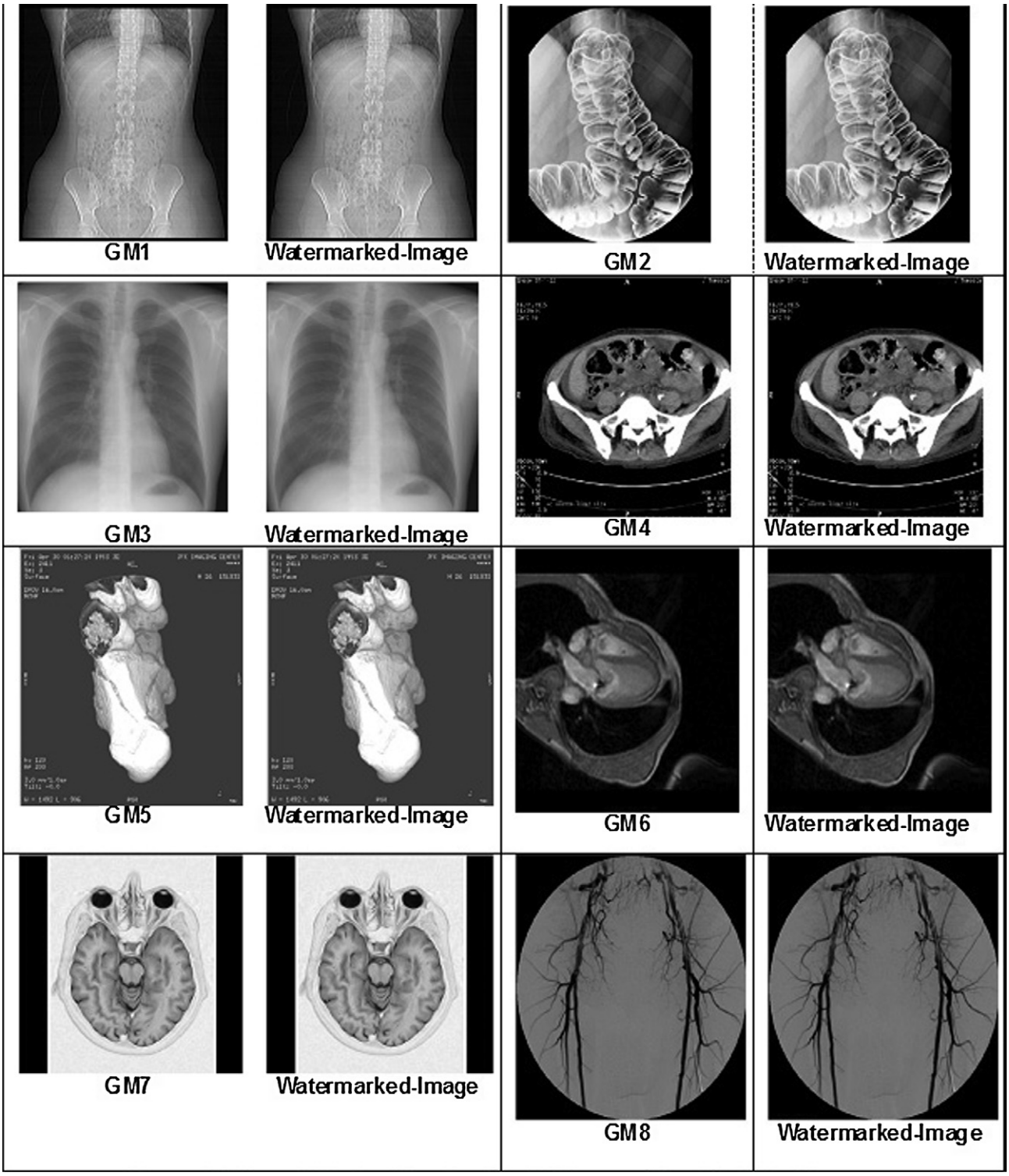

Figure 5 The original and authentecated gray seal medical test images.

Table 2 Rtsuto of Color Medical.

Nuu	BER %	SNR(dB)	MSE	PSNR (dB)
CM1	0	48.06	0.124	57.20
CM2	0	57.33	0.097	58.26
CM3	0	55.90	0.077	59.27
CM4	0	55.64	0.0128	59.56
CM5	0	53.35	0.128	57.06
CM6	0	45.37	0.13	56.99
CM7	0	46.76	0.175	55.70
CM8	0	48.27	0.133	56.89
Average	0	51.33	0.117	57.62
Standard deviation	0	4.72	0.034	1.31

supports the connection of networked printers, such as laser imagers.

4.1. DICOM file structure

A DICOM file contains the header and the image data [22]; the header could be the patient's name, type of scan, image dimensions, etc., the header and the image data are stored in the same file. DICOM file requires a 128-byte preamble, followed by the letters 'D', 'I', 'C', 'M', then followed by the data set of the image (Fig. 3). The data elements are partitioned into logical groups to describe the image attributes, patient information, hospital information and doctor's information. The Data Element structure (Fig. 4) and consist of the following

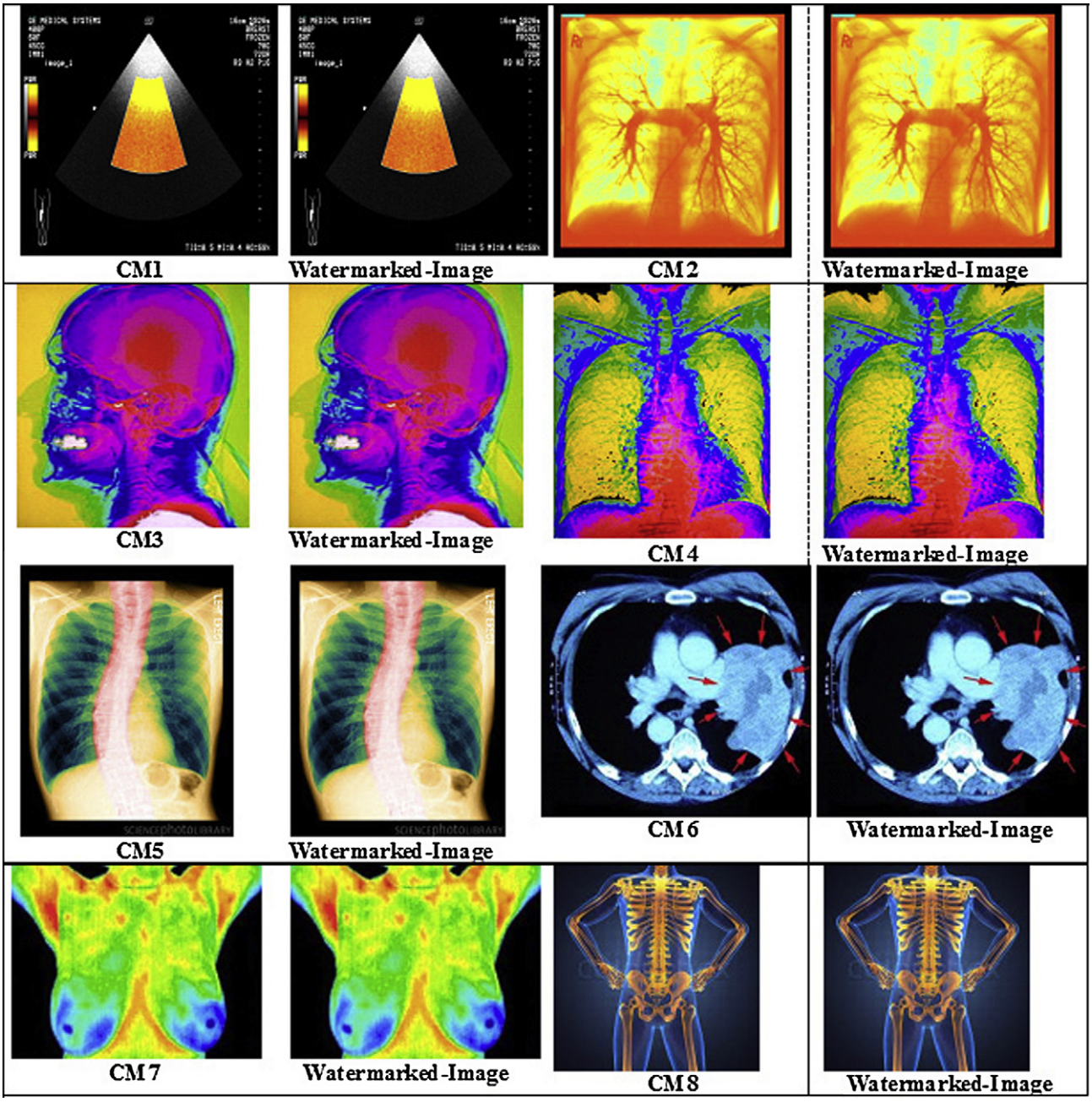

Figure 6 The original and authentecated color medical test images.

fields: Tag, Value Representation, Value Length and the Value.

- Tag: The identifier of the data element; it consists of 32 bits unsigned integer, 16 bits for the Group Number, and 16 bits of the Element Number.
- Value Representation: It specifies the data type of the value field (byte, integer, character).
- Value Length: It specifies the length of the value field (number of bytes).
- Value: It represents the data value of this data element.

4.2. Authentication of the DICOM file

DICOM file can be authenticated using the proposed technique. Before applying the technique the pixel values must be extracted from the DICOM file as follows:

1. Load the PIXEL-DATA data element buffer.
2. Load the needed data elements to extract the pixel values' matrix such as:
   - Bit allocated: Number of bits for each pixel.
   - Bits Stored: Number of the bits which actually used to describe the pixel value.
   - Pixel Representation: it determines if the value is saved as an unsigned integer or 2s complement number.
   - Transformation Parameters: They are Window Center and Window Width.
3. Extract only the bits used for describing the pixel.
4. There may be a transformation can be used to retrieve the real pixel values; this is used to appear hidden details in the image for specific use of this image, and the transformation parameters are embedded in the DICOM file.
5. After having the image matrix, the R–S-Vector, Hash Value can be determined using the proposed technique, and they can be embedded in the image matrix, and then the image matrix is rewritten to the file according to the DICOM standard.

To check the authentication of DICOM file, first the image matrix must be extracted, and from it; the R–S-Vector and the embedded hash value are extracted; the hash value at the receiver are determined then the two hash values are compared to check the integrity of the image.

5. Experimental results

The experimental results of the proposed technique for authentication and integrity of medical images based on reversible watermarking technique are discussed in this section. An application is created using C# language [23] and MATLAB application [24] to implement this technique. The main application and user interfaces are programmed using C# language, and they call MATLAB functions to read and write DICOM images, and to compress the R–S-Vector, in addition to calculate the performance parameters. The performance parameters that are determined to measure the performance of the proposed technique are: Signal to Noise Ratio (SNR), Mean Square Error (MSE), and Bit Error Rate (BER).

This section consists of four subsections: the first subsection represents the performance results due to applying the proposed technique on grayscale DICOM images. The second subsection represents the results of applying the technique on color medical images. The third section represents the results of applying the proposed technique on grayscale test images and on the color test images.

5.1. Experimental results of grayscale medical images

The proposed technique is applied on eight grayscale DICOM medical images (GM1, GM2, ..., GM8); the results are presented in Table 1, the original images (before embedding a watermark), and the authenticated images (after embedding the watermark) are shown in Fig. 5. Using the human visual system to detect differences between the original images and the related authenticated images (Fig. 5), clearly no one can detect any difference between them; therefore, the proposed system can embed the watermark inside the original grayscale DICOM medical images without any noticeable distortion. The experimental results (Table 1) show that the BER is equal to zero for all eight images, this means that the bit stream sequence (watermark) extracted from the image at the receiver side is identical to the bit stream sequence embedded in the image at the sender side.

These results prove that the proposed technique is totally revertible, and the original images can be retrieved at the receiver side without any distortion because of the R–S-Vector is extracted without errors, in addition to the ability of assuring the integrity of the images and the authentication of the sender because of the hash function can be extracted without errors. To check the ability of the technique to discover if the integrity and authentication of the image are corrupted, at the sender side a one pixel of the authenticated image was modified, then at the receiver side the proposed technique detected this modification and sent a message that the image is not longer authenticated.

From Table 1, the minimum SNR is 49 dB with average equal about 56.63 dB and standard deviation 11.29; It, consequently, indicates that the SNR has great values; therefore, corruption due to embedding the watermark in the original image is very low. The maximum value of MSE is 0.156 with average equal about 0.124 and standard deviation 0.0.02; It indicates that the MSE has very low consistent value; therefore, the embedded watermark does not affect the quality of the original images. The PSNR has large consistent values; the minimum value equals 56.2 dB while the high value equals

Table 3 Results of grayscale test images.

Name	BER %	SNR (dB)	MSE	PSNR (dB)
Gl	0	50.10	0.121	57.30
G2	0	50.35	0.154	56.26
G3	0	53.96	0.137	56.76
G4	0	51.05	0.153	56.28
G5	0	48.12	0.108	57.80
G6	0	47.50	0.143	56.58
G7	0	51.30	0.157	56.17
G8	0	51.56	0.128	57.06
Average	0	50.49	0.138	56.78
Standard deviation	0	2.03	0.018	0.58

58.13 dB with the average equals 57.25 dB, and standard deviation equals 0.7.

Experimental results are similar to experimental results of grayscale medical image (Section 5.1); therefore, it indicates that: the corruption due to embedding the watermark in the original image is very low; the embedded watermark does not affect the quality of the original images, and the identical original image can be extracted at the receiver side.

5.2. Experimental results of color medical images

The proposed technique is applied on eight color (RGB) medical images (CM1, CM2, ..., CM8); the results are presented in Table 2, the original images and the authenticated images are shown in Fig. 6. Using the human visual system to detect differences between the original images and the related authenticated images (Fig. 6), clearly no one can detect any difference between them; therefore, this technique can embed the watermark inside the original color medical images without any noticeable distortion. The experimental results (Table 2) show similar results as in Table 1, the BER is zero for all six color images. The minimum SNR range starts from 45.37 dB with an average about 51.33 dB and standard deviation 4.72. The average MSE is about 0.117 with 0.034 standard deviation. The PSNR has large consistent values; the minimum value equals 56.89 dB while the maximum value equals 59.56 dB with the average equals 57.62 dB, and standard deviation equals 1.31.

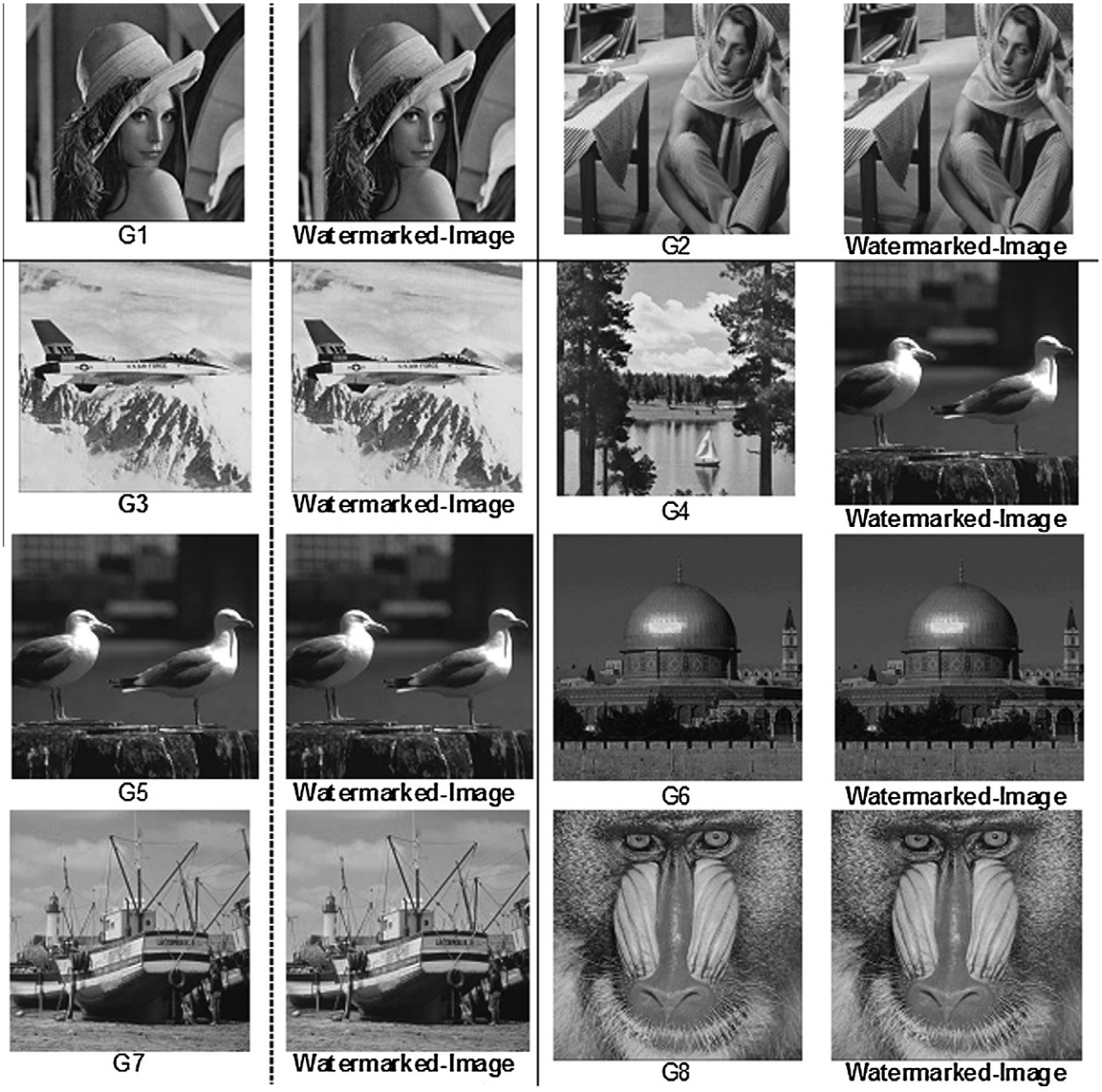

Figure 7 The original and authentecated color medical test images.

Table 4 Results of color test images.

Name	BER%	SNR (dB)	MSE	PSNR (dE)
Cl	0	54.09	0.135	56.83
C2	0	54.25	0.09	58.59
C3	0	51.43	0.103	58.00
C4	0	53.76	0.089	58.64
C5	0	53.48	0.149	56.40
C6	0	47.4S	0.129	57.02
C7	0	50.24	0.116	57.49
C8	0	50.10	0.154	56.26
Average	0	51.85	0.121	57.40
Standard deviation	0	2.45	0.025	0.93

5.3. Experimental results of grayscale test images

The proposed technique is applied on eight grayscale test images (G1, G2, . . . , G8). The results are presented in Table 3, the original images and the authenticated images are shown in Fig. 7. And it is applied on eight color test images (C1, C2, . . . , C8). The results are presented in Table 4, the original images and the authenticated images are shown in Fig. 8.

Using the human visual system to detect differences between the original images and the related watermarked images for both of grayscale and color images (Figs. 7 and 8), Clearly no one can detect any difference between them. The experimental results for grayscale and color images (Table 3 and 4) are similar to the results of grayscale and color medical images (Tables 1 and 2). Table 3 shows that the BER is zero for all eight images. The minimum SNR is 47.50 dB with an average

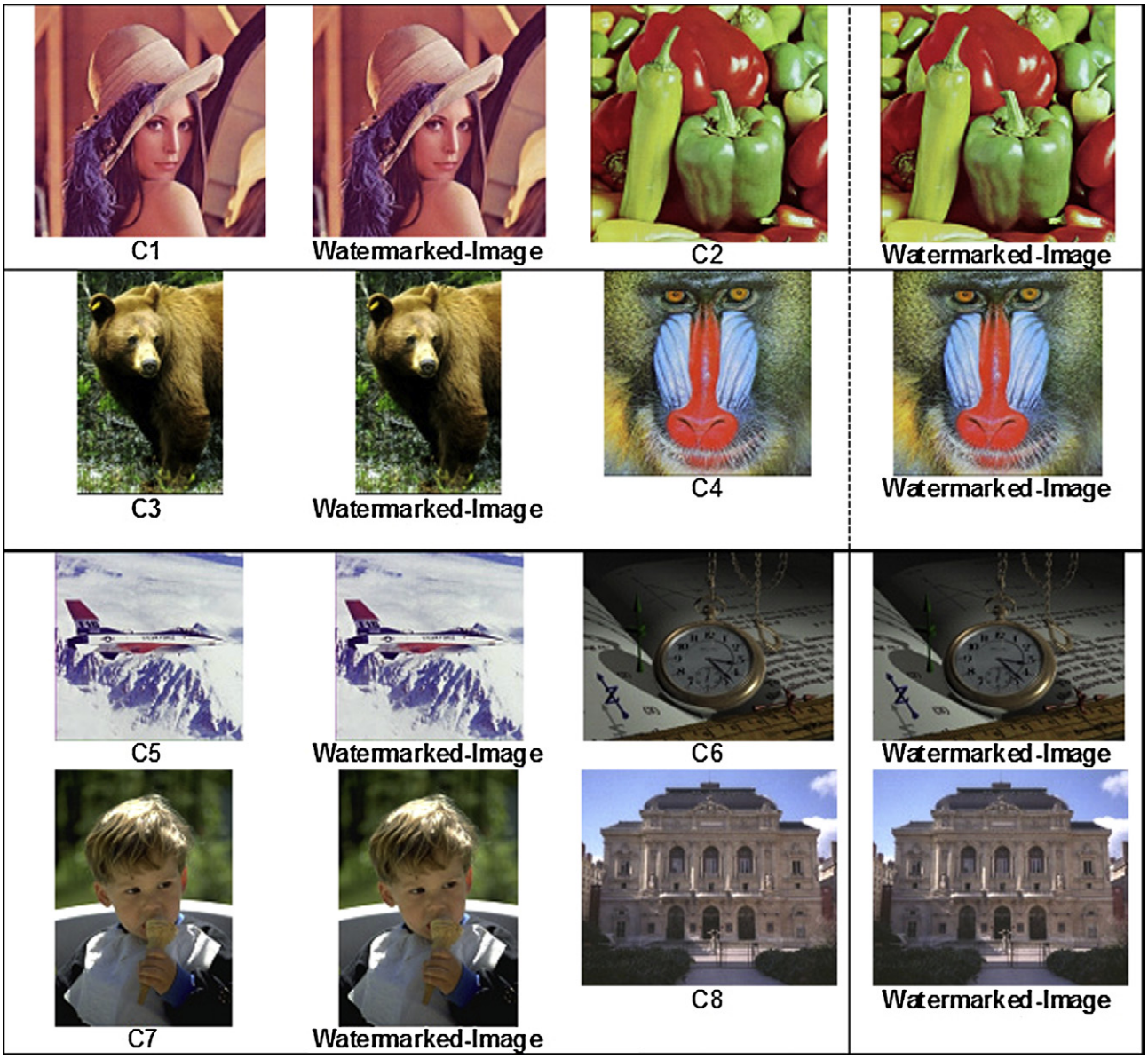

Figure 8 The original and authentecated color medical test images.

about 50.49 dB and standard deviation 2.03. The minimum PSNR equals 56.17 dB, and the average equals 56.78 dB, and the standard deviation equals 0.58. The MSE is about 0.138 on average with standard deviation 0.018. And the experimental results for color images (Table 4) show that the BER is equal to zero for all eight images, the minimum SNR is 47.48 dB with an average about 51.85 dB and standard deviation 2.03, and the average MSE equal is 0.121 with standard deviation 0.025. Clearly, the performance measurements of the grayscale and color images are consistent with the performance measurements of the DICOM gray medical image and the color medical images.

The proposed technique is not designed to provide copyright protection; however, it is proposed to provide integrity and authentication services for the DICOM images. Therefore, its target is not to be robust against modification attacks, but its target is to detect any modification into the watermarked images. To check the ability of this technique to discover if the integrity and authentication of the image are corrupted, at the sender side a one pixel of the authenticated image was modified, then at the receiver side the proposed technique detected this modification and sent a message that the image is not validated. These checks were implemented for all types of test images that are used in section four.

6. Conclusion

In this paper, a DICOM image security technique based on the reversible watermarking method is proposed, this technique provides system authentication service, image integrity service and patient information confidentiality service; it is reversible because the original medical image can be retrieved at the receiver side without any distortion.

At the sender side to achieve integrity service a hash value based on MD5 is determined from the image. And to satisfy reversible feature R–S-Vector is determined and is compressed based on Huffman's compression algorithm, then the compressed R–S-Vector. The hash value and the patient-ID are concatenated to produce a watermark, and to provide confidentiality and to protect hash value. The watermark is encrypted using AES encryption technique.

The proposed technique is implemented using C# language and MATLAB application, and its performance is tested using eight grayscale DICOM images, eight color medical images, eight grayscale test images and eight color test images. The experimental results prove that the proposed technique can provide system authentication service, image integrity service and patient information confidentiality service with high efficiency. Practically, this technique provides adequate embedded capacity (around 0.3 bpp) to hide the proposed watermark.

Concluded results show that the BER equal 0 for both of grayscale DICOM, color medical images, grayscale and color test images. Consequently, the bit stream sequence (watermark) extracted from the image at the receiver side is identical to the bit stream sequence embedded in the image at the sender side, and these results prove that the proposed technique is totally revertible.

The mean SNR equals about 52 dB with standard deviation 3 for all images (color and grayscale), this indicates that the SNR has consistent large values; this proves that the corruption due to embedding the watermark in the original image is very low. And the mean MSE equals about 0.12 with standard deviation 0.02; this indicates that the MSE has very low consistent values. Hence the embedded watermarks do not affect the quality of the original images. The PSNR has large consistent values; the average PSNR equals about 57 dB with a standard deviation equals about 0.93; therefore, the proposed technique creates watermarked images that have high imperceptibility, invisibility and transparency.

As a future work the proposed technique can practically be included within the medical information systems to provide medical image integrity, system authentication and confidentiality. Other revertible watermarking methods can be proposed to increase the amount of embedded data, and other lossless compression methods can be proposed to enhance the ability of the proposed technique to embed larger amount of data. In addition, a public encryption technique can be used instead of AES encryption technique to provide source authentication.

References

[1] Coatrieux G, Montagner J, Huang H, Roux C. Mixed reversible and RONI watermarking for medical image reliability protection. In: 29th International conference of the IEEE, engineering in medicine and biology society, Lyon, 2007. p. 5653–6.

[2] Memon N, Gilani S. Adaptive data hiding scheme for medical images using integer wavelet transform. In: IEEE international conference on emerging technologies, Islamabad, Pakistan; 2009. p. 221–4.

[3] Lim Y, Feng D. Multiple block based authentication watermarking for distribution of medical images. In: International symposium on intelligent multimedia, video and speech processing, ISIMP 2004, Hong Kong; 2004. p. 631–4.

[4] Boucherkha S, Benmohamed M. A lossless watermarking based authentication system for medical images. Int J Signal Process 2004;1(4):278–81.

[5] Mostafa S, El- sheimy N, Tolba A, Abdelkader F, Elhindy H. Wavelet packets-based blind watermarking for medical image management. Open Biomed Eng J 2010;4:93–8, <http://www.benthamscience.com/open/tobej/openaccess2.htm>.

[6] Memon N. Watermarking of medical images for content authentication and copyright protection. PhD thesis, Pakistan: Faculty of Computer Science and Engineering, GIK Institute of Engineering Sciences and Technology; May 2010.

[7] An L, Gao X, Xuelong L, Tao D, Deng C, Li J. Robust reversible watermarking via clustering and enhanced pixel-wise masking. IEEE Trans Image Process 2012;21(8):3598–611.

[8] Das S, Kundu M. Effective management of medical information through a novel blind watermarking technique. J Med Syst 2012;36(5):3339–51.

[9] Pan W, Coatrieux G, Cuppens N, Cuppens F, Roux C. Reversible watermarking based on invariant image classification and dynamical error histogram shifting. In: Engineering in medicine and biology society, annual international conference of the IEEE, August, 2011. p. 4477–80.

[10] Ko L, Chen J, Shieh Y, Hsin H, Sung T. Nested quantization index modulation for reversible watermarking and its application to healthcare information management systems. Comput Math Method Med 2012;2012:1–8, <http://www.hindawi.com/journals/cmmm/2012/839161/> [Article ID 839161].

[11] Bouslimi D, Coatrieux G, Roux C. A joint encryption/watermarking algorithm for verifying the reliability of medical images: application to echographic images. Comput Methods Programs Biomed 2012;106(1):47–54.

[12] Rahimi F, Rabbani H. A dual adaptive watermarking scheme in contourlet domain for DICOM images. Biomed Eng Online

2011;10:1–18, <http://www.biomedical-engineering-online.com/content/10/1/53>.
[13] Lee H, Rhee K. High Capacity Lossless Data Hiding. In: Second international conference on information systems security, ICISS 2006. Kolkata, India: Springer; 2006. p. 326–36.
[14] Shi YQ. Reversible data hiding. In: International workshop on digital watermarking 2004. Seoul: Lecture Notes in Computer Science 3304; 2004. p. 1–13.
[15] Goljan M, Fridrich J, Du R. Distortion-free data embedding. In: Proceedings of the 4th information hiding workshop, Pittsburgh; April 2001. p. 27–41.
[16] Stallings W. Cryptography and network security: principles and practice, 5/E. Prentice Hall; 2011. [ISBN-10:0136024858, ISBN-13:9780136024859].
[17] The MD5 Message-Digest Algorithm. RFC 1321 – IETF, April 1992. <http://www.ietf.org/rfc/rfc1321.txt>.
[18] Advanced Encryption Standard (AES). FIPS 197, November 26, 2001.
[19] Sayood K. Lossless compression handbook. Academic Press; January 3, 2003. [ISBN-10:0126208611, ISBN-13:978-0126208610].
[20] Sayood K. Introduction to data compression. 3rd ed. Morgan Kaufmann; 2005 [ISBN-10:012620862X, ISBN-13:978-0126208627].
[21] Salomon D. Data compression: the complete reference. 4th ed. Springer; 2006 [ISBN-10:1846286026, ISBN-13:978-1846286025].
[22] Digital imaging and communications in medicine (DICOM), Part 5: Data structures and encoding. Med Imag Technol Alliance, 2012. <http://medical.nema.org/>.
[23] Microsoft Visual Studio 2008. Express editions; 2012. <http://www.microsoft.com/visualstudio/en-us/products/2008-editions/express>.
[24] MATLAB. The language of technical computing. MathWorks Company; 2012. <http://www.mathworks.com/products/matlab/index.html>.