International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 8, No. 1 (2015), pp. 49-60
http://dx.doi.org/10.14257/ijsip.2015.8.1.05

High Capacity, Reversible Data Hiding Using CDCS Along with Medical Image Authentication

Amrinder Singh Brar[1] and Mandeep Kaur[2]

[1,2]*Department of Information Technology*
Panjab University, Chandigarh, 160014, India
reachamrinder@gmail.com, mandeep@pu.ac.in

Abstract

Healthcare institution that handles a number of patients, opinions is often sought from different experts. It demands the exchange of the medical history of the patient among the experts which includes the medical images, prescriptions and electronic patient records (EPR) etc. In order to reduce storage and transmission cost, data hiding techniques are used to embed patient information with medical images. In medical imaging applications, there are stringent constraints on image fidelity that strictly prohibit any permanent image distortion by the watermarking or data hiding. Authenticity is another important aspect in medical image watermarking. This paper proposed modified difference expansion watermarking using LSB replacement in the difference of virtual border for data hiding in medical images. The Class Dependent Coding Scheme (CDCS) is used to encode the EPR data so that embedding capacity can be increased. The image hash is calculated using MD5 to provide authentication. Experimental results show that proposed scheme provide us large data hiding capacities along with very high PSNR values as compared to earlier EPR data hiding techniques.

Keywords: *EPR, Reversible Watermarking, Virtual Border, CDCS, LSB replacement, Data hiding, Difference expansion*

1. Introduction

Telemedicine combines Medical Information System with Information Technology that includes use of computers to receive, store and distribute medical information over long distances. Telemedicine can be divided into number of medical related technologies using computers for healthcare like teleradiology, telepathy, telecare, telesurgery, teleneurology *etc.*, [11, 1]. In number of medical applications, medical images require special safety and confidentiality, because critical judgment is done on the information provided by medical images. Critically ill or injured patients can be treated locally by effective and secured communication between remote hospitals and distant specialist [7]. Exchange of medical images between hospitals located in different geographical locations is a common practice now a day as shown in, Figure 1. Hence, healthcare industry demands secure and more information hiding techniques promising strict secured authentication and communication through internet or mobile phones. Digital image watermarking provides copyright protection to digital image by hiding appropriate information in original image to declare rightful ownership [2].

ISSN: 2005-4254 IJSIP

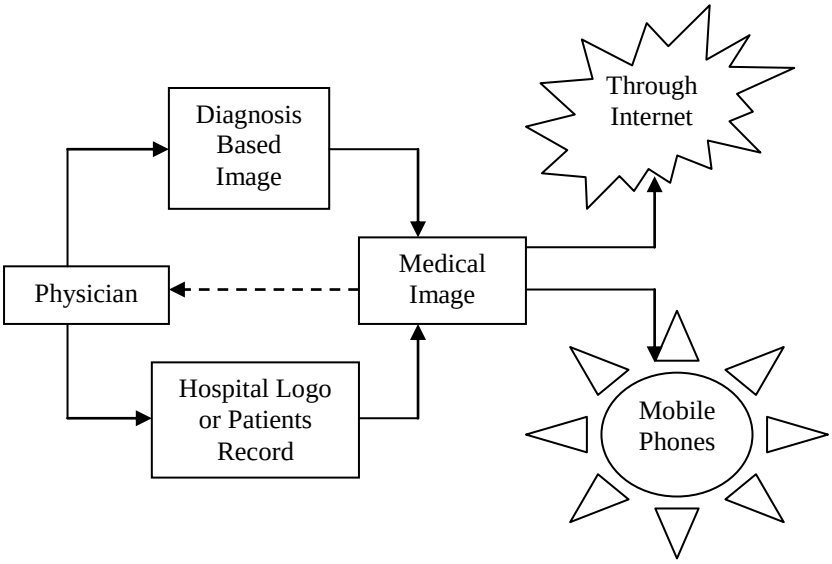

Figure 1. Transmission of Medical Images from Physician of Remote Hospital to Specialist Through Internet and Mobile Phone

All works reported in data hiding in medical image are watermarking for authentication and EPR hiding. The medical images of different modalities with EPR attached to them can be sent to the clinicians residing at any corner of the globe for the diagnosis. Embedding of EPR with medical images will save storage space of the Hospital Information System, enhance confidentiality of the patient data and save the bandwidth required for transmission. Obviously this will reduce the cost of diagnosis. This kind of a system requires a high level of security, which can be ensured by using digital watermarking techniques. Hiding patient data in the medical image is one of the applications of digital image watermarking. The patient data in the electronic format is called Electronic patient record (EPR) [3].

In quality-Sensitive applications such as medical imaging, military imaging, law enforcement, and remote sensing where a slight modification can lead to significant difference in final decision making process, the original image without any modification is required during image analysis. Even if the modification is quite small and imperceptible to human eyes, they do not accept because it may affect the right decision and lead to legal problems. In multimedia archives, content providers do not want to waste their storage keeping both the original image and the watermarked one, due to cost and maintenance problems [4]. Consequently, reversible data hiding techniques are designed to solve the problem of lossless embedding of large messages in digital images so that after the embedded message is extracted, the image can be completely restored to its original state before embedding occurred [5].

Application wise, robust watermarks are suitable for copyright protection because they can resist common image processing operations. On the other hand, fragile watermarks can be used to detect tampering and authenticate an image because it is sensitive to changes. Semi-fragile watermarks are usually applied in some special cases of authentication and tamper detection. These cases may consider lossy image compression as legitimate changes while highlighting geometrical distortions as intentional attacks [10].

The Proposed technique is developed based on the works implemented in the year 2010 [8, 9]. Nova Hadi Lestriandoko and Taufiq Wirahman [9] developed exciting reversible watermarking technique using difference of virtual border for digital image protection. This technique provides very high PSNR values as compared to any reversible

watermarking technique developed earlier. Due to very good PSNR values this technique proves to be best for medical image watermarking. On the other hand Sunita V. Dhavalen and L. M. Patnaik [8] developed excellent CDCS encoding scheme for encoding EPR data and then used statistical quantity for embedding EPR data in RONI of the medical image. In this technique the focus is purely on increasing robustness so PSNR values are not as good as provided by virtual border embedding technique. The CDCS used in this embedding scheme proves to be very efficient encoding scheme due to its less complexity and very good PBS (Percentage Bit Saving) [8] values. Even though techniques like Huffman encoding provides higher PBS (Percentage Bit Saving) than CDCS, but due to its higher complexities Huffman encoding is not considered to be one of the useful encoding technique for such applications and is problem for those who have less efficiency in understanding the logics of such complex encoding techniques, they need simple and efficient encoding schemes. So we used CDCS to encode EPR data and embed encoded EPR data along with medical image hash into the cover image using difference of virtual borders. This scheme provides us very efficient, high capacity, reversible data hiding scheme for medical images along with image authentication.

1.1. The Rest of the Paper is Organized as Follows

In Section 2, proposed scheme is discussed in detail. The discussion is done on data embedding as well as data extraction techniques to be used. In Section 3, Experimental results are displayed and short discussion is written on the basis of these results. Section 4 provides the conclusion of the paper by discussing the artifacts of proposed technique. The future scope of this proposed technique is discussed in Section 5. Last but not the least Section 6 provides the detail of references.

2. Proposed Scheme

This scheme works in two phases: CDCS phase [8] and the virtual border embedding phase [9]. In CDCS phase stream of EPR encoded bits along with hash code bits are prepared and in virtual border embedding phase difference of virtual border is used to embed the FBS (Final Bit Stream) into the original medical image.

2.1. CDCS

In the proposed scheme the CDCS is used to encode the EPR data. CDCS assigns fixed codes to each character by considering their probability of occurrences as shown in, Figure 2 [8] In this encoding scheme the EPR data is divided into three non-overlapping classes according to their occurrence frequencies. Class A consists of most frequently appearing character set, Class-B consists of average frequently occurring character set and Class-C consists of less frequently appearing character set.

This scheme is prepared to encode only capital letters, alphanumeric and few special characters so four bits are needed to represent each character in respective class as shown in, Table 1. Variable length class codes are prepared using Huffman encoding as given in Table 2. So any character can be represented by only 4 bits prefixed by 1-bit or 2-bit class code [6].

Therefore, Class code along with character code can distinguish 48 different characters [8] These 48 characters are sufficient to represent any EPR.

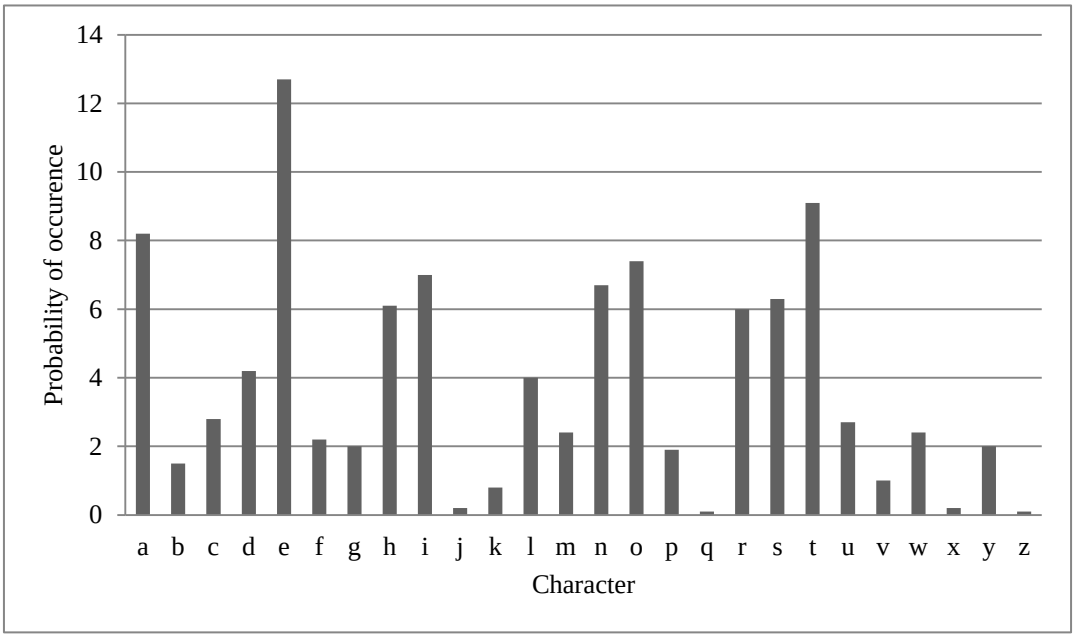

Figure 2. Probability of Occurrences for EPR Characters [8]

Table 1. Fixed Codes of each Character within Class [8]

Class A	Class B	Class C	4-Bit Code
Blank	M	0	0000
.	U	1	0001
E	G	2	0010
T	Y	3	0011
A	P	4	0100
O	W	5	0101
N	B	6	0110
R	V	7	0111
I	K	8	1000
S	X	9	1001
H	J	(	1010
D	Q	)	1011
L	Z	=	1100
F	,	*	1101
C	-	%	1110
:	_	+	1111

2.2. Efficiency of CDCS

Huffman encoding provides variable length encoding and is very complex encoding scheme. Such complex encoding schemes proves to be very inefficient in the EPR encoding, where doctors have to encode EPR data again and again without having deep

technical knowledge about such complex encoding scheme. On the other hand ASCII gives fixed length codes for the characters [9]. The CDCS combines the advantages of both fixed length and variable length coding to represent each character with less no of bits.

2.3. Percentage Bit Saving

If Z_1, Z_2 and Z_3 are the total number of characters belonging to Class-A, Class-B and Class-C respectively,

Total number of bits to be embedded is given by,

$m=(Z_1 + 2Z_2 + 2Z_3) + 4h$

Where, $h=Z_1+Z_2+Z_3$, i.e. total number of characters in EPR file.

So Percentage Bit Saving (PBS) [11] is given by

$$PBS = \left[1 - \left(\frac{m}{7h}\right)\right] \times 100\%$$

Table 2. Class Codes for Three Classes

Class	Class code	Length
A	1	1-bit
B	00	2-bits
C	01	2-bits

2.4. Hashing Original Image

After EPR data bits are encoded, Original medical image will be hashed using MD5 hash algorithm. The MD5 Message-Digest Algorithm [8] is a cryptographic hash function that produces a 128-bit (16-byte) hash value. Then it is converted into binary and appended with EPR data bits to get the Final Bit Stream (FBS) which is to be embedded in original medical image.

2.4. Virtual Border Embedding Phase

After the Final Bit stream is prepared, Difference of Virtual border is used to embed the FBS into the original image. In virtual Border embedding technique we use the difference of horizontal virtual border to embed Prepared Final Bit Stream [9]. Virtual border is mirror of image border line.

First step of watermarking is to create virtual border. The horizontal borders are copied to obtain the mirror of first and last row. This increases the image size. The difference of these mirrors are used to embed an image signature. The Final Bit Stream (FBS) is embedded using LSB in the difference of virtual border [9]. The flow chart of proposed embedding scheme is shown in Figure 3.

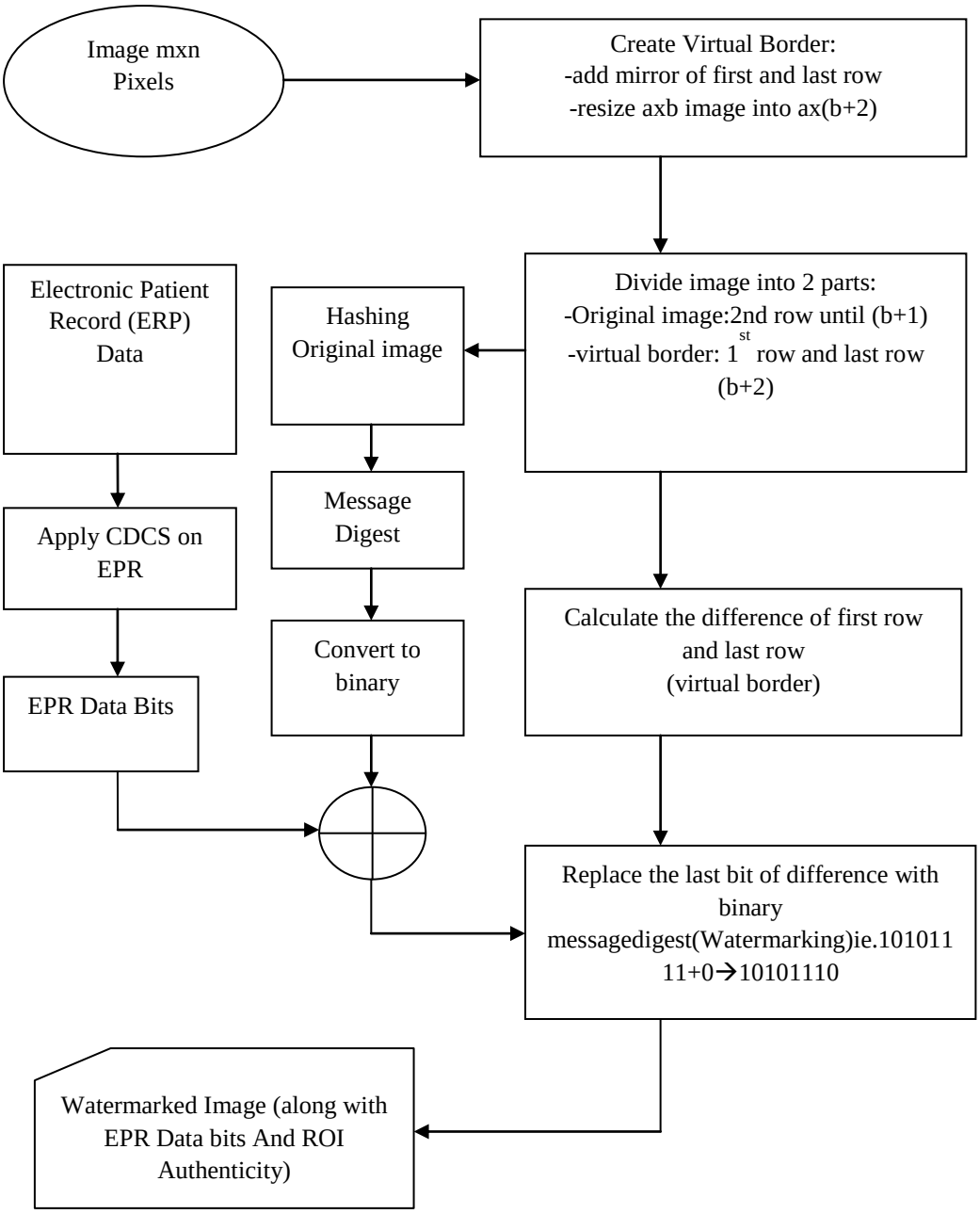

Figure 3. Flowchart of Proposed Embedding Scheme

In this technique, the payload is embedded in the difference of virtual border. For a pair of pixels values (x, y) in a RGB image (24 bits), $x, y \in Z$, $0 \leq x_{RGB}, y_{RGB} \leq 255$, we can define the difference h as

$$h = | x - y |$$

If we assumed that the higher value is called *top* and the smaller is called *bottom* , so we have

$h = top - bottom$, where $top \geq bottom$

To embed the payload, we use LSB (Least Significant Bit) method and replace the last bit of difference with payload. It makes the chance of new difference only have two possibilities: equal or changes For example: we have a difference 10100101, it will be embedded by a payload (the payload are "0" or "1"). The chances of new difference are

"10100100" or "10100101". So, the range of new difference h' is $0 \leq h \leq 255$. The values of h' is defined as

h+1 →last bit=0, payload=1
h → last bit=payload
h-1 →last bit=1, payload=0
The new pixel values can be calculated as (x′, y′)
(x+1,y) →x≠255, h′=h+1
(x, y-1) →x=255, h′=h+1
(x-1,y) →h′=h-1
(x,y) →h′=h

2.5. EPR Data Retrieval and Authentication.

For EPR data retrieval and authentication of medical image we can extract the embedded bit stream from the LSB of horizontal border differences. This is done as follows:

1. Calculate the difference of horizontal border line and get the Final Bit Stream (FBS) which was embedded. Now in this FBS first 128 bits are image hash which will be used to authenticate the image. Rest are EPR data bits.
2. Erase the virtual border and hash image using MD5. This will reduce the image size.
3. Authenticate the medical image by comparing the image hash in step1 and step2, if authentication succeeds, restore the original image and go to next step otherwise drop the image as unauthenticated.
4. Apply inverse CDCS on the EPR bits received in step1 to get the embedded EPR data.
5. Original image is restored without virtual borders and the data is stored in .xls format file.

The flow chart of data retrieval and authentication phase is displayed in Figure 4.

3. Experimental Results and Discussion

Experiments are carried out on certain categories of medical images varying in size and bits per pixels. Standard results are taken using various medical images. To evaluate the data hiding capacity and medical image quality, the performance evaluation is measured by Percentage Bit Saving (PBS), Peak Signal to Noise Ratio (PSNR), and Mean Square Error (MSE). Various sets of experiments are carried out to see the effect of CDCS and Difference of Virtual Border embedding schemes on PSNR.

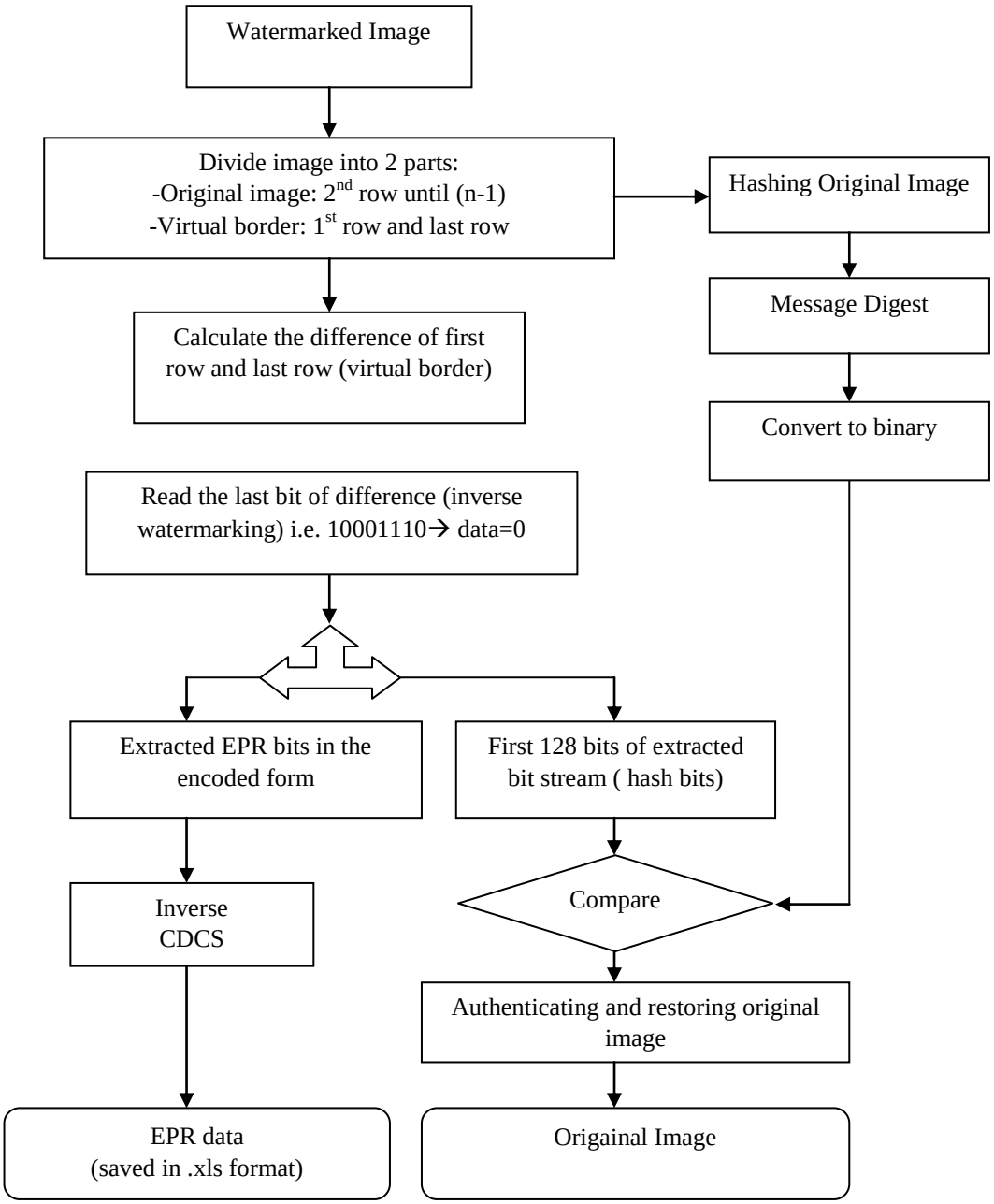

Figure 4. Flow Chart of Data Retrieval and Authentication Phase

3.1. Performance Evaluation of Proposed Technique

Quality after extraction is of outmost importance in medical image watermarking. The evaluations of results for proposed scheme are stored in Table. 3. This evaluation is done using standard gray scale radiological medical image files (BMP and JPEG) of size 512×512, having 8 bits per pixel. The PSNR is calculated after embedding EPR files of various sizes *i.e.*, containing 183 characters, 270 characters and 343 characters. The graphical results shown in Figure 5, had drawn the conclusion that proposed technique provides us high PSNR values which is not achieved by some other data hiding techniques. Figure 6 shows the graphical representation of PSNR and MSE values of the four medical images in table below, this graphical analyses is done using Graphical User Interface prepared for proposed

scheme. GUI allows users (doctors, lawyers, surgeons *etc.*,) to interact with proposed scheme in much simpler way. Figure 7 displays GUI developed for proposed scheme. The snapshots of results and GUI (Figure 6, Figure 7) is taken during implementation of proposed scheme using MATLAB 2007b.

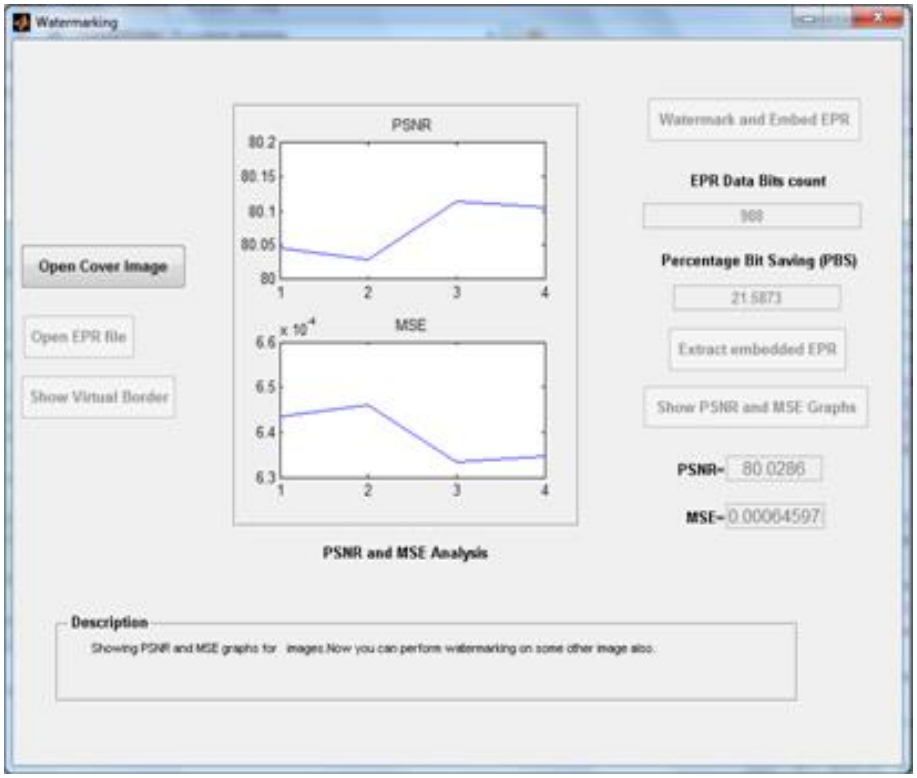

Figure 6. Graphical Analyses of Proposed Scheme in GUI taking 240 EPR Characters and Set of Four Medical Images

Table 3. PSNR Values of Images Using Proposed Scheme

Medical Image Under Test	PSNR (db) with Proposed Scheme when EPR contains 183 Characters	PSNR (db) with Proposed Scheme when EPR contains 270 Characters	PSNR (db) with Proposed Scheme when EPR contains 343 Characters
L-Spine	82.00	79.72	78.93
MRI head	82.46	80.02	79.15
Abdomen	82.34	80.04	79.17
Shoulder	82.54	80.11	79.19

3.2. Performance Analyses of Proposed Technique for some other Images

The efficiency of proposed technique is also experimentally audited using RGB and gray scale medical images of 24 bit per pixel (bpp). Excellent performance is achieved using proposed technique. The results are represented in, Table 4 for various medical images. This analyses proves the performance of proposed scheme for medical images of varying sizes and types.

Table 4. Performance Analysis of Proposed Scheme by Taking Images of Various Sizes

S. No	Medical Images	Size of Original Image	Size of Water--marked Image	PSNR (db)	MSE	PBS
1	Bone	960x971	960x973	85.08	0.00020	21.58
2	Petct	995x984	995x986	85.65	0.00017	21.58
3	Abnl	995x859	995x861	84.58	0.00022	21.58
4	Saddle	1053x876	1053x878	84.90	0.00021	21.58
5	Keosys	1440x900	1440x902	86.50	0.00014	21.58
6	Brain	432x538	432x538	78.52	0.00075	21.58
7	Mri	2000x2687	2000x2689	92.76	0.00000034	21.58
8	Fetal	927x608	927x610	82.86	0.00033	21.58
9	Head	640x496	640x498	80.66	0.00055	21.58

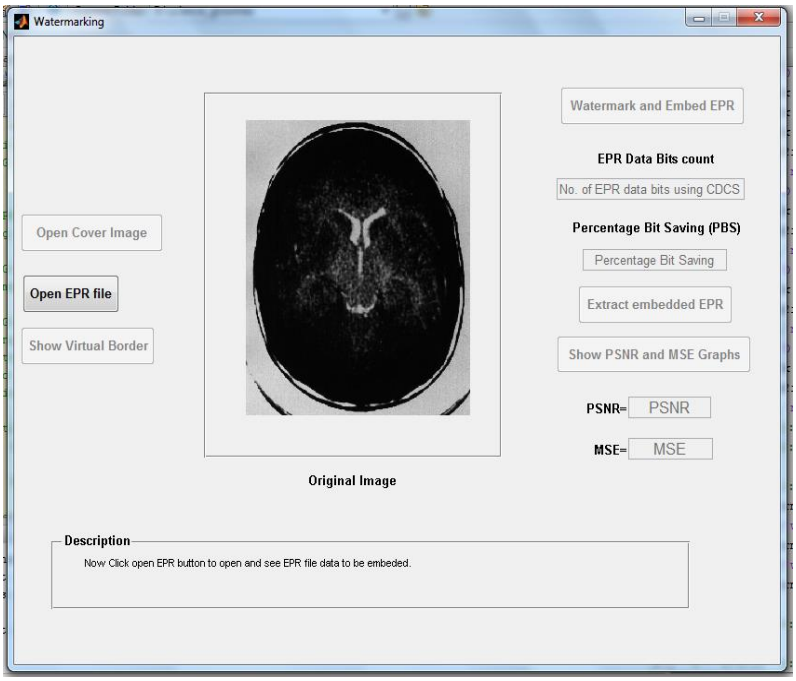

Figure 7. Display of Graphical User interface (GUI) of Proposed Scheme

4. Conclusion

The proposed scheme proves to be efficient reversible data hiding technique along with medical image authentication. Experimental results shows that proposed scheme provides high data hiding capacity along with very high PSNR values, also this scheme provides lossless EPR data hiding. We used CDCS to prepare EPR data and

watermarking(embedding) is done using Pixel Difference of Virtual Borders. Image Hash using MD5 algorithm is used to provide authenticity to medical images. Virtual Border embedding provides us high PSNR values as compared to other techniques(statistical quantity 'α' embedding scheme) when used to embed EPR data in the medical images. Thus proposed technique can be effectively used for embedding high volume EPR data in medical images with very low image distortions.

The only limitation of proposed technique is change of image size, although it is small (only increase 2 rows). But the increase in storage capacity is very minute and almost negligible as compared to high PSNR value provided by this technique. So proposed techniques proves to be best for environments where one need to embed large EPR data and any change in medical image is not acceptable, so that medical image can be recovered without any degradation.

5. Future Scope

For future work, the technique can be extended to allow any tamper detected in ROI region of medical image to be recovered; this can be done by embedding ROI pixels along with payload in the medical image.

References

[1] D. Osborne, "Embedded Watermarking for Image Verification in Telemedicine", PhD Thesis, University of Adelaide, **(2005)**.
[2] B. L. Gunjal and S. N. Mali, "ROI Based Embedded Watermarking of Medical Images for Secured Communication in Telemedicine", International Journal of Computer and Communication Engineering, **(2012)**.
[3] K. A. Navas, M. Sasikumar and S. Sreevidya, "A Benchmark for Medical Image Watermarking", Transaction on Google.
[4] S. Lee, C. D. Yoo and T. Kalker, "Reversible image watermarking based on integer-to-integer wavelet transform", IEEE Transc, on Information Forensics and Security, vol. 2, no. 3, **(2007)**, pp. 321-330.
[5] S. L. V. Krishna, B. Abdul Rahim, F. Shaik and K. SoundaraRajan, "Lossless Embedding using Pixel Differences and Histogram Shifting Technique", IEEE, **(2010)**, pp. 213-216.
[6] S. N. Mali and S. V. Dhavale, "High Capacity Secured Adaptive EPR Data Hiding with Integrity Checking using CDCS", International Journal of computational sciences, vol. 3, no. 6, **(2009)** December, pp. 657-668.
[7] D. Osborne, D. Rogers, J. Mazumdar, R. Coutts and D. Abbott, "An Overview of Wavelets for Image Processing for Wireless Applications", Proceedings of SPIE: Smart Structures, Devices and Systems, University of Melbourne, Australia, vol. 4935, pp. 427-435.
[8] S. V. Dhavale and L. M. Patnaik, "High Capacity, Robust Lossless ERP Data hiding using CDCS with ROI Tamper Detection", Int'l conf on computer & communication Technology |ICCCT'10|, IEEE, **(2010)**, pp. 108-112.
[9] N. HadiLestriandoko and T. Wirahman, "Reversible Watermarking Using Difference of Virtual Border for Digital Image Protection", 2010 International Conference on Distributed Frameworks for Multimedia Applications(DFmA), **(2010)**.
[10] C.-S. Woo, "Digital Image Watermarking Methods for Copyright Protection and Authentication", Theses submitted in accordance with the regulations for Degree of Doctor of Philosophy at Information Security Institute, Queensland University of Technology, **(2007)**.
[11] N. Ahmed Memon, "Watermarking of Medical Images for Content Authentication and Copyright Protection", PhD Thesis, GIK Institute of Engineering Sciences and Technology Topi, Swabi, Khyber Pakhtunkhwa, Pakistan, **(2010)** May.

Authors

Er. Amrinder Singh Brar, recieved the B.Tech. degree in Computer Science and Engineering from Punjab Technical University, Jalandhar, India in 2010. Soon after his graduation, he joined Master of Engineering in Information Technology from Panjab University, Chandigarh, India. He did his specialization in Medical Image Watermarking. From 2012-2014, he joined as an Assistant Professor in Department of Computer Science and Engineering at BFCET, Bathinda, India. He was awarded two times with Annual Academic Excellence Award during his job. He is currently pursuing regular Ph.D. degree from Punjabi University, patiala, India. His major research areas are Image Processing, Steganography and Digital Watermarking, Data Hiding. He is also focusing his recent research on Digital Holography.

Er. Mandeep Kaur, she recieved her B.Tech degree from Punjab Technical University, Jalandhar, India in 1999. After her graduation she worked as a Lecturer in various reputed engineering colleges. Then she completed her Master of Engineering degree from Punjab Engineering College (PEC), Chandigarh, India in 2004, and started working as a lecturer in PEC, chandigarh, India. After one year she joined Department of Information Technology, Panjab University, Chandigarh, India as an Assistant Professor. She is also pursuing her Ph.D. degree from Panjab University, Chandigarh. She works in the field of Digital Image Processing.

