Analysis of Medical Images by Data Hiding Techniques

Kriti
Student, M.Tech
ECE Dept., BGIET, Sangrur

Shalvi
Assistant Professor
ECE Dept., BGIET, Sangrur

ABSTRACT

A fragile watermarking system is designed for the authentication of medical images. The proposed system authenticates the computed tomography (CT) scan medical images of the effected organs against different distortions. The system enhances the embedding capacity of a CT scan image by isolating the region of interest (ROI) based on higher gray values which need to be unaffected for medical diagnosis and hide character watermark only in region of non interest (RONI) and border areas, thus not compromising the diagnostic value of medical imaging. The method utilizes the spatial domain overlapping and least significant bit (LSB) replacement method (in case of overflow or no data matched) for embedding the watermark. Experimental results reveal that the proposed system detects both legitimate and illegitimate distortions and outperforms the existing reversible data hiding schemes in terms of embedded capacity and PSNR.

Keywords

Digital watermarking, ROI, RONI, Data authentication, CT scan image, Embedding, Patient report.

1. INTRODUCTION

In this modern era of technology medical images can be given to the patient directly or send to the patient by online and also maintained as a soft and hard copy at the hospital for diagnosing and later in the future purposes. The problem arises here is that while sending or giving the data to the patient it has to be found whether the data belongs to particular patient or not and also the privacy of the patient is a primary concern [1]. Hence authentication is required. If we consider medical images especially ultrasound and MRI, they are a confidential property of the patients or defense personals and need to be authenticated and transmitted without any vulnerable attack called tempering. However it’s not possible in this hacking area, so we need to provide some security especially in the region of interest to avoid manipulations which define the defected area of the patient. A number of methods are emerging and watermarking is one of them. Some people take watermarking as steganography but there is a difference between the two. Watermarking is defined as the practice of imperceptibly altering a work to embed a message about that work. On the other hand, steganography is the practice of undetectable altering a work to embed a secret message [2]. Digital image watermarking has gained a great interest in the last twenty years among scholars who provide a number of methods. However, still we are far away from being fully or accurately successful. Therefore, more and more people are entering the field to make the watermarking idea useful and reliable for the digital world. Of these various watermarking methods, some beat others in terms of basic watermarking requirements like robustness, invisibility, processing, cost etc.

The whole system of watermarking comprises of an adder called Embedder, which embeds the mark and a detector as shown in the figure below. The Embedder’s output is typically transmitted or recorded. At the detector side, the embedded image is accessed as an input to the detector. If a payload is present, detector detects and extracts the encrypted data and payload. Watermarking, like cryptography, needs secret keys to identify legal owners. The key is used to embed the watermark, and at the same time to extract or detect it. The embedded signal can be exposed only with a right key. While a single bit of information indicating that a given document is watermarked or not is sufficient sometimes, most applications demand extra information to be hidden in the original data. This information may consist of ownership identifiers, transaction dates, serial numbers etc. That plays a key role when illegal providers are being tracked.

In this paper, a reversible ROI-based watermarking scheme is proposed which is capable of hiding patient's data and verifying authenticity of ROI. In section 2, watermarking techniques for medical images are reviewed. In section 3, the proposed watermarking technique, including data hiding and extracting on the receivers end is presented. In section 4, experimental results are provided to demonstrate the efficiency of the scheme. Finally, in section 5 the conclusion is drawn.

2. LITERATURE REVIEW

The earlier watermarking techniques were proposed for data hiding applications. These days, the authentication ability became an important factor in medical image watermarking techniques. Few of them are discussed below:

Pushpala, K., Nigudkar., R., [1] proposed a novel watermarking technique for medical image authentication. In this paper, medical image was stored in PACS (picture archiving and communication system) that were accessed over the intranet by radiologists for diagnosis .This system used for security measures in the information system of the hospital to ensure integrity of medical image data that was being transferred over the public network.

Lee, H.K., Kim, H.J., Kwon, K.R., and Lee, J.K., [2] proposed ROI medical image watermarking using DWT and Bit-plane. In this paper, digital watermarking and bit-plane used for embedded data into medical image. They purposed the technique of embedding the ROI image into the spatial domain of non-ROI part of a medical image.

Raul, C.R., Claudia, F.U., Gershom, T.B, [3] proposed data hiding scheme for medical image. In this paper, digital image watermarking had been proposed as a method to enhance medical data security, confidentiality and integrity.

Weng, S., Zhao, Y.,Pan, J.S and Ni, R., [4] proposed Reversible data hiding using the companding technique and improved DE Method .In this paper , the watermark was embedded into high frequency sub bands of the integer wavelet transform (IWT), using the compounding technique. The result showed high visual quality in moderate capacities.

Huang, H.C., B fang, W.C. and Chen, S.C., [5] proposed Privacy protection and authentication for medical image with record-based watermarking .This paper presented a practical scheme for privacy protection and authentication of medical image with the aid of EXIF metadata and associated records of the patients. This application was robust and alters selected coefficient in the transform domain to accomplish the embedding process.

Khurshid, K., Faure, C. and Vincent, N., [6] proposed Fusion of word spotting information for figure Caption retrieval in historical document image. This paper presented a method for figure caption detection by employing a fusion of several information sources. The evaluation was performed on document gathered from the collection of the historical medical digital library medic.

Sun, X. and Bo, S., [7] proposed a blind digital watermarking for colour medical image based on PCA .This paper presented Robust digital image blind watermark scheme that was used to protect colour medical image. In this approach, K-L transform was applied to an RGB medical image and the binary watermark was embedded into low frequency sub-band of DWT of the principal component of medical image.

Saravanan, P. and Nagarajan, S., [8] proposed an adaptive learning approach for tracking data using visual and features. This paper proposed the concept of image and text association, a cornerstone of cross media web information fusion. Two learning methods for discovering the underlying association between image and texts based on small training data sets were proposed.

Poonkuuntran, S. and Rajesh, R.S., [9] proposed. A messy watermarking for medical image authentication. In this paper, the digital funds image was chosen for simulation and analysis of the proposed scheme. These images were given in TIF format in RGB colour space. The proposed scheme dynamically generates the watermarking using messy models. It was embedded inside the image by expanding intra plane difference between any two colour planes of image.

Proposed Technique: Our proposed authentication watermarking approach is based both on the fragility of embedding mechanism and on matching the embedded features with the features extracted from a test image. The detection of tampering relies on both the embedding mechanism and the embedded data. The alteration can also be located unless there is global tampering or the tampering involves large areas. We present our approach in the context of grayscale images with tiff compression. A block diagram of the embedding process is given in Fig 1. Watermarks are inserted as an overlap if pixel values matched or divided into two different parts and located or inserted in two locations if not matched to the required pixel value decide by ASCII values into the quantized coefficients via using ASCII values of the data to be hidden in a sequence concatenated together. Explained below are two aspects of watermark-based authentication, namely, to embed what data and how to embed them.

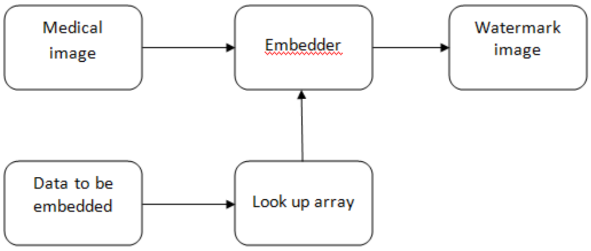

Fig 1: Embed which data and how to embed

3. METHODOLOGY

3.1 Generating the watermark

In order to generate the watermark, following steps are implemented:

1. Read the data from the word file which need to be hidden and for watermarking. In our case we use patient report. Read the text file containing the patient information; convert each character of text file into its corresponding ASCII code.

2. Now concatenate the data into a single line array "Combine" having length say N, such that LOOK UP ARRAY=Combine = $\{combine\ (i) \in [0, 1], 1 \leq i \leq N\}$.

3. Convert each ASCII code into its corresponding binary code and form the vector which may have length of M bits such that vector = $\{w2(i)|w2(i) \in [0, 1], 1 \leq i \leq M\}$.Note that the length of W2 depends upon the number of characters used to represent Electronic Patient Record (EPR) multiplied by 8, as 8-bit binary representation is used to represent ASCII codes into the binary form. In our simulations, 9872 characters of data are used, therefore the value of M is 9872 × 8 = 78976 bits.

3.2 Embedding data into medical image

The embedding process starts with the generation of watermark. Later on the watermark is embedded in RONI.

The process is described step by step as follows:

1. Generate the watermark and put ASCII values of concatenated data in array "combine".
2. Generate an empty array for secret keys.
3. Put MAC in secret key
4. Put the pixels of test images in an array.
5. Generate a dummy array called LOOK UP ARRAY in order to put the ASCII values of Concatenated data "Combine"
6. Generate an empty array "values= [ ]" which is needed in the scanning process in order to keep temporary data used for decision making.
7. Scanning of the test image from left to right row wise in order to find a corresponding match for pixel values from the dummy array.
8. If matched pixel finds in test image first check the secret key array if that pixel area has been used in secret key array or not.

9. If yes, scan for another location.

10. If no, put the values of that row and column number in the secret key array.

11. If space, put the pixel value in the border area by making sure that the location should not repeat in a secret key.

A) If it does not find similar value pixel scanning while i.e. all pixels are used to choose the pixel which has nearest value.

B) Update the encrypted image array according to this newly found pixel. And update he secret key.

Resulted image has now become encrypted or embedded images which need to be transmitted.

3.3 Extraction process

Since proposed scheme is blind so there is no need of original image to extract the embedded watermark.

The extraction process has the following steps:

1. Read the encrypted image.
2. Load the secret key produced in the embedded process. It needs to be send separately along with encrypted image.
3. Extract the pixels by using the secret key.
4. Decrypt the extracted watermark and MAC by converting back to characters in order to make string.
5. Compare the extracted MAC (hash value) to the computed hash. If both are same, received image is authentic, otherwise declare it as unauthentic.
6. Put the decoded data in a txt. File.

4. RESULTS

4.1 Experimental results

Original Dimensions of brain and lung image are [263 260] and [187 174] respectively which have been reshaped to different scales in order to get the results. Equations used for evaluating PSNR and MSE are

PSNR = 10log10 *(R^2/MSE)

MSE = I, J[test image(i, j) – embedded image(i, j)]2/ j

where R is the maximum fluctuation of intensity in the input image data type.

The degradation in terms of PSNR and MSE in the test image and watermarked image for different images are shown in Table 1, 2 and 3.

Tables show the degradation in visual quality of the watermarked image with respect to the original image by embedding watermarks of varying strengths in terms of PSNR and MSE.

4.2 Experimental data and results

Fig 2 shows the test images of brain and lung respectively.

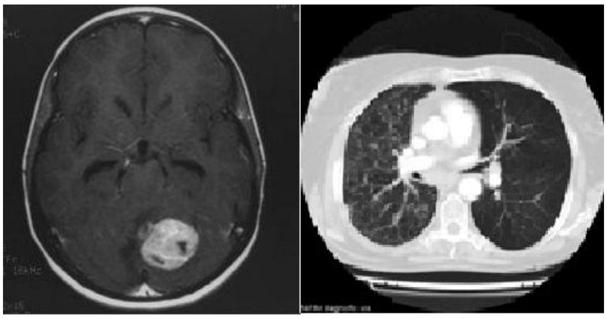

Fig 2: brain and Lung CT Scan images

Table 1. For brain CT scan image at constant scaling factor

Image	**Payload (bytes)**	**MSE**	**PSNR (dB)**	**Scaling factor**
Brain image (263*260)	11k	0.0475	61.3657	3
	9.8k	0.0252	64.1254	3
	6.8k	0.0033	72.8838	3
	4.9k	0.0009	78.5718	3

Table 2. For brain CT scan image at constant payload

Image	**Payload (bytes)**	**MSE**	**PSNR (dB)**	**Scaling factor**
Scan (263*260)	9.8k	0.0623	60.19	2.8
	9.8k	0.0252	64.1254	3
	9.8k	0.0034	72.7759	3.5
	9.8k	0.0009	78.30	4

Table 3. For image of lung organ at diffenrt payload and scaling factor

Image	**Payload (bytes)**	**MSE**	**PSNR (dB)**	**Scaling factor**
Brain image (263*260)	9.8k	0.0089	68.6345	4
	6.8k	0.0017	75.9308	4
	9.8k	0.2754	53.7308	3
	6.8k	0.0287	68.5538	3

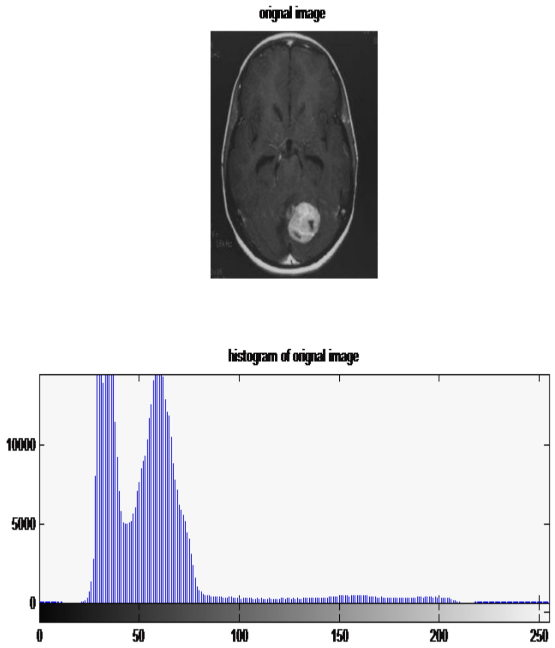

Fig 3: Original test image and histogram with payload 9.8k bytes

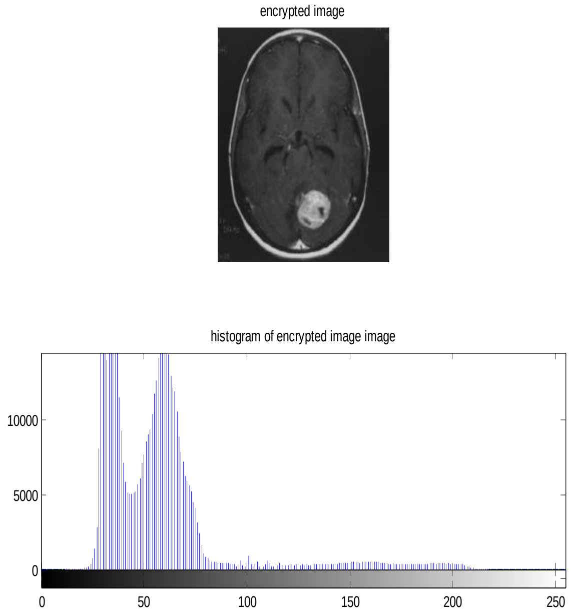

Fig 4: Encrypted image and histogram with payload 9.8k bytes

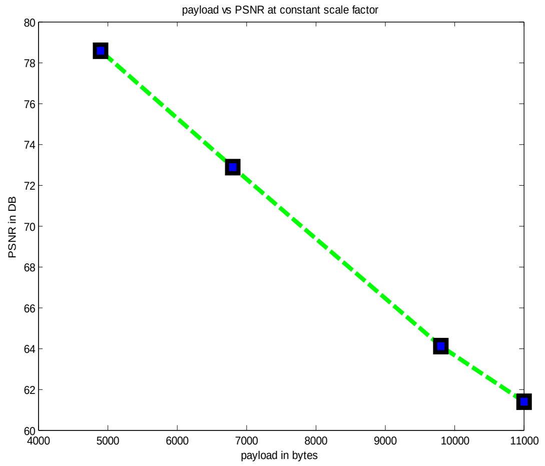

Fig 5: Payload vs psnr for brain image at constant scaling factor

It is found out that the visibility of watermark reduced when little payload has been embedded and increased as we move on to increase the payload bits.

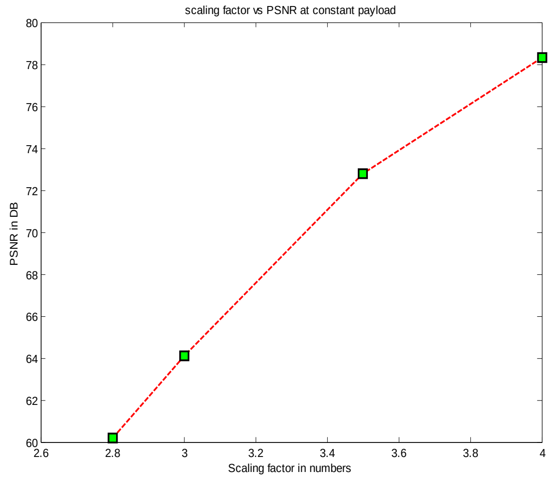

Fig 6: PSNR vs scaling factor at constant payload (brain image)

It is found that when we keep payload constant Visibility of watermark decreases with increasing scaling factor.

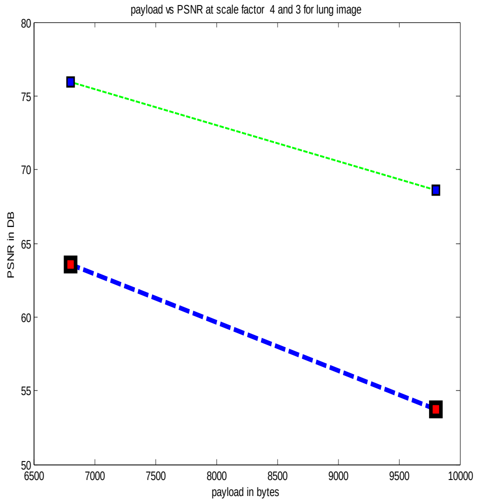

Fig 7: Green line for 4 scaling factor, blue line for 3 scaling factor (LUNG IMAGE)

Above figure is a plot between payload and PSNR values for the lung image. They can be analyzed in the same way as is done for brain image.

It is hard to locate the difference in the intensity bins of original and encrypted histograms as the changes are minute but can be seen in there difference images.

Authentication: The integrity of images at the receiving end is checked by comparing the decoded string which comes as output in extraction process. If it is unaltered i.e. it matched with the information send separately. Then It is authenticated. Otherwise it has been tampered in the transmitted media. Gaussian noise, median filtering, compression are some of the attacks which can result in tampering of image and hence we can't authenticate the image if tampered.

Computational efficiency has been increased by this method as the algorithm needs to scan only once per character to embed the data in the image. It took 176.299045 seconds to embed data in brain CT scan image when rescaled with factor 3 and to embed approx 10 kilobytes of payload.

5. CONCLUSION AND FUTURE SCOPE

The fragile data hiding technique is proposed in the spatial domain to preserve the history of medical image by embedding the medical diagnosis report and other data. While embedding the data, ROI of medical image is avoided to ensure the integrity of ROI. The scheme allows the simultaneous storage and transmission of electronic patient record along with image authentication information which can be extracted at the receiving end without the original image. Encryption of the embedded data is done to provide additional security but fails to authenticate if tampered by different means it also provides sufficient capacity for storing about more than 10k of data when rescaled. The scheme can easily be used in e-diagnosis applications but still needs to improve the algorithm in the process of authentication.

6. REFERENCES

[1] Pushpala, K. and Nigudkar, R., 2005. "A novel watermarking technique for medical image authentication", IEEE, pp. 683-686.

[2] Lee, H.K., Kim, .H.J., Kwon K.R. and Lee .J.K., 2005. "Digital watermarking of medical image using ROI information'', IEEE, pp. 404 -407.

[3] Raul, C.R. and Claudia, U.F, 2007. "Data Hiding Scheme for Medical Images", IEEE, pp. 32-36.

[4] Weng, S., Zhao, Y., Pan, J.S., and Ni, R., 2007. "A novel high capacity reversible watermarking scheme," IEEE International Conference on Multimedia and Expo (ICME 07), vol. 3, pp. 723 – 730.

[5] Huang, A.H.V.C., Fang, B.W.C. and Chen, S.C., 2009. "Privacy Protection and Authentication for Medical Images with Record-Based Watermarking", IEEE, pp. 190 – 193.

[6] Khurshid, K., Faure, C. and Vincent, N., 2009. "Fusion of word spotting and spatial information for figure caption retrieval in historical document image'', IEEE, pp. 266-270.

[7] Sun, X. and Bo, S., 2010. "A Blind Digital Watermarking for Color Medical Images Based on PCA", IEEE, pp. 421 – 427.

[8] Saravanan, P. and Nagarajan, S., 2010. "An Adaptive Learning Approach for Tracking Data using Visual and Features", IEEE, pp. 192 – 196.

[9] Poonkuntran, S. and Rajesh, R.S., 2011. "A Messy Watermarking for Medical Image Authentication", IEEE, pp. 418 – 422.