Text Fusion in Medical Images Using Fuzzy Logic Based Matrix Scanning Algorithm

Pinki Jain, Anu Aggarwal (Guide)

Doon Valley Institute of Engineering and Technology, Kurukshetra University, Karnal, India
Faculty of Doon Valley Institute of Engineering and Technology, Kurukshetra University, Karnal, India

Abstract- Text fusion in medical images is an important technology for image processing. We have lots of important information related to the patient's reports and need lots of space to store and the proper position and name which relates that image with that data.

In our work we are going to find out the AOI (area of interest) for the particular image and will fuse the related document in the NAOI (non area of interest) of the image, till yet we have many techniques to fuse text data in the medical images one of form them is to fuse data at the boarders of the images and build the particular and pre defined boarder space.

We are going to propose an algorithm called fuzzy logic based matrix scanning algorithm in which we will first find out the area of interest and after that we find noisy pixels of the image to embed data in that noisy portions to save the boarder size. Our proposed technique is LSB to store text data in pixels. We use **MATLAB** for carrying out implementation on our proposed work.

Index Terms- Electronic Patient Record, medical images, Text data, text fusion

I. Introduction

Due to the development of latest technologies in communication and computer networks, exchange of medical images between hospitals has become a usual practice now days. 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 clinical images, prescriptions, initial diagnosis etc. With the increasing use of Internet, these digital images can be easily accessed and manipulated. Considering patient's privacy and diagnostic accuracy, the prevention of medical images from tampering tends to be an urgent task. It is required to imbibe the aspects of confidentiality, authentication and integrity with the distribution of these images in the Health Information System. Medical images are exchanged for number of reasons, for example teleconferences among clinicians, interdisciplinary exchange between radiologists for consultative purposes, and distant learning of medical personnel. Most hospitals and health care systems involve a large amount of data storage and transmission such as administrative documents, patient information, and medical images, and graphs. Among these data, the patient information and medical images need to be organized in an appropriate manner in order to facilitate using and retrieving such data and to avoid mishandling and loss of data. In order to overcome the capacity problem and to reduce storage and transmission cost, data hiding techniques are used for concealing patient information with medical images. Those data hiding techniques can be also used for authentication. These applications demand large amount of patient information available in one single image rather than over several entities.

In medical images, AOI is an area which contains important information and must be stored without any distortion. In this paper , we present an fuzzy logic based matrix scanning algorithm which finds the noisy pixels in the medical images by scanning the whole image in several directions using 3*3 scanning window .Then patient data is hidden inside these pixel.

A. Related Researches

1. In 2005, Hung kyo lee proposed digital watermarking technique for medical image that prevents illegal forgery that can be caused after transmitting medical image data remotely[1]. A wrong diagnosis may be occurred if the watermark is embedded into the whole area of image. Therefore, they embed the watermark into some area of medical image, except the decision area that makes a diagnosis so called region of interest (ROI) area in our paper, to increase invisibility. The watermark is the value of bit-plane in wavelet transform of the decision area for certification method of integrity verification. The experimental results show that the watermark embedded by the proposed algorithm can survive successfully in image processing operations such as JPEG lossy compression.

2. In 2007, Rodríguez-Colín Raúl, Feregrino-Uribe Claudia, Trinidad-Blas Gershom de J. proposed watermarking scheme that combines data compression, encryption and watermarking techniques and image moment theory applied to radiological medical images[2]. In this work we use DICOM data as a watermark to embed in medical images. Image quality is measured with metrics which are used in image processing such as PSNR and MSE. Our results show good accuracy in the watermark extraction process.

3. In 2009, Liu Xin a semi-fragile digital watermarking algorithm[3] is proposed based on integer wavelet transform to estimate integrity and authenticity of medical images. Using matrix norm quantization, it embeds watermarks into medium-frequency an high-frequency detail sub-bands of medical images' integer wavelet domain and it is a blind-testing watermarking algorithm. Attacking experiments show that the algorithm not only has both robustness and sensitivity, but also exactly locates distorted area, so it is an effective semi-fragile digital watermarking algorithm for medical image field. Semi-fragile

watermarks require not only certain robustness but also sensitivity, so this paper proposes a semi-fragile watermarking algorithm based on integer wavelet transform and matrix norm quantization, in which watermarks are separately embedded into medium-frequency and high frequency in medical images' integer wavelet domain.

II. PROPOSED METHOD

Our proposed work consists of two modules:

A. Module 1

It is the first module we read the host image file on which we are going to implement our algorithm to MATLAB workspace and do the needful changes on that like if the image is in RGB format we convert it into grayscale and if the size of the image is too big then convert it to the nominal size.Then the normalization of light intensity for the image is to be done so that we can get more specific intensities to work with that.After the normalization we proceed with the finding the noisy pixels in the image by various scanning methods. We are going to implement a new fuzzy logic based (5x5) and (3x3) matrix scanning algorithm which will scan the whole picture in several directions and according to the direction of the majority carries the direction on the preceding pixel will be decided. According to which we can able to decide whether the point is a noise or not.

B. Module 2

After detecting the noisy coordinates of the image we add our data to them as a watermarking content and save the positions where we are going to fuse our patient data in image. After fusing the data in noisy coordinates we again calculates the reaming data to be hide or not and if there will be any data left to fuse in the image then our algorithm proceed towards hiding the data in the image boarder and generates only that much of boarder which is needed to be store the remaining data and cuts rest of the non area of interest to save image.

C. Fusion of patient data in medical image

- Fuzzy logic based matrix scanning algorithm

As we propose an algorithm which finds the non area of interest in the medical image. A 3*3 scanning window is used to find the non area of interest in medical image. Our algorithm works as follows:

1. Initialize Naoi (non area of interest) equal to zero.
2. Then, separate the color planes of the image .We will embed the patient data in any one color plane of the image. We separate green color plane and starts work on this.
3. The size of the resulting image is determined.
4. Our algorithm completes when
   The total length of the data > Naoi
5. For j = 1 to col -2
   For i= 1 to row – 2
   Where i,j are the elements of the scanning window.and row and col are the total no of rows and columns of the window.
6. Initialize votes = 0;
   Where votes are the predictions of finding noisy pixels by the neigbours.
7. For m= 0 to 2
   For n= 0 to 2
   Subtract the image (i+1,j+1) from the image (i+m,i+n)
   if its subtraction is greater than the threshold value then

   otes = votes+1;
   end
   end
   end
8. If(votes>4)
9. Naoi _add=[naoi_add;i+1 j+1];
   Naoi=naoi+1;
   end
   end
   end
   th = th -2;
   end
10. Noisy pixels computed.

Hence noisy pixels computed.we hide the patient data into these pixels .our hiding technique is LSB (Least significant bit).

D. Methodology

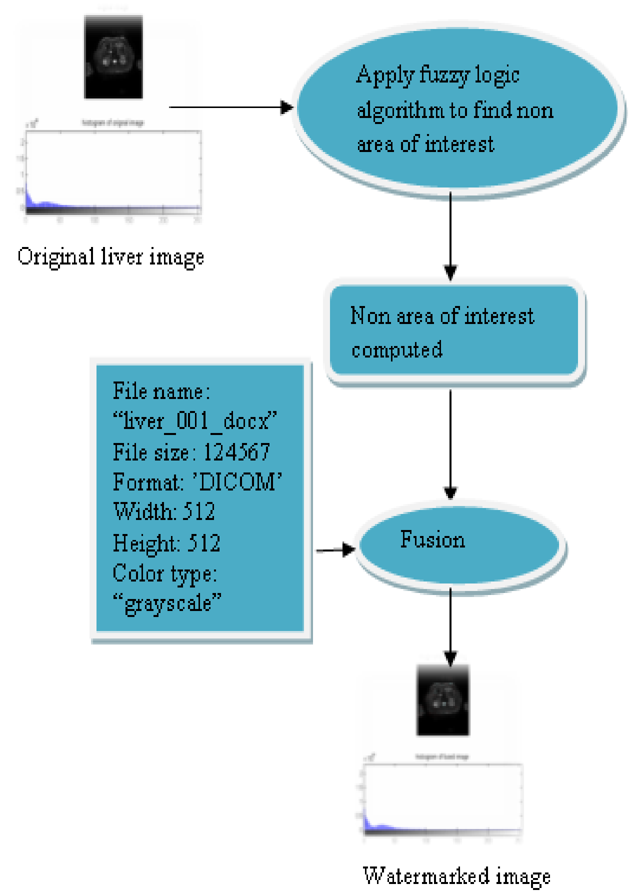


Fig 1: Fusion process of patient data in medical image

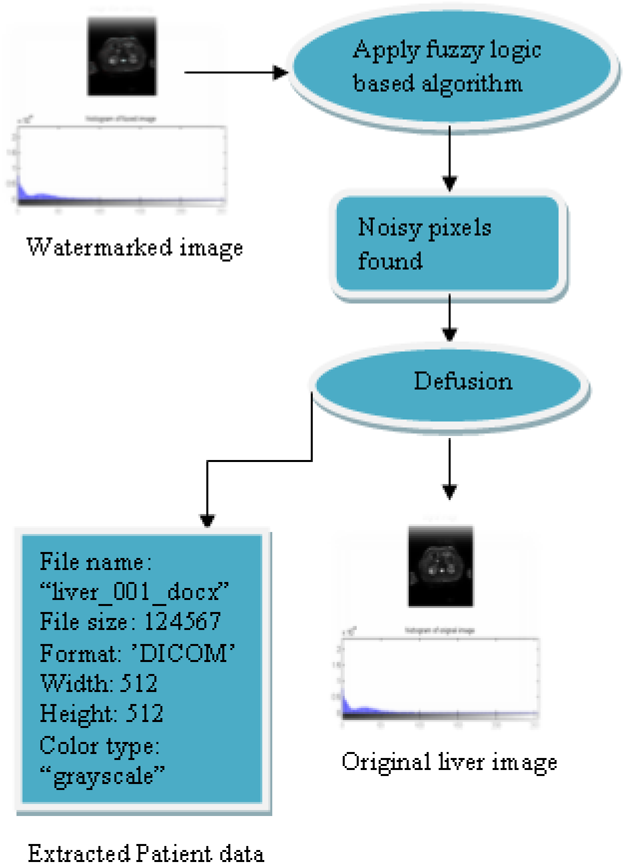


Fig 2: Defusion of patient data from the watermarked image

E. Defusion of patient data from the fused image

This is the reverse process of fusion of patient data in medical image .

1) we finds the noisy pixels in watermarked image.

2) after finding the non area of interest, we can extract the patient data from the image.

III. RESULTS

In this section, we present the results obtained with our work .To evaluate our proposed algorithm, we simulated with the medical images; Ctscan, Brain and liver with the size of 512*512 that converted patient files of DICOM to image data using MATLAB 7.7.0(R2008b). In our scheme ,the extracted watermark is a visually recognizable pattern. The viewer can compare the results with referenced watermark subjectively. However ,.the subjective measurement is dependent on factors such as the expertise of the viewers and the experimental conditions. Therefore. a quantitative measurement is needed to provide objective judgment of the extracting fidelity. We define the similarity measurement between the original image and the defused image by the two quality metrices.

1. PSNR(power signal-to-noise ratio)
2. MSE is mean square error

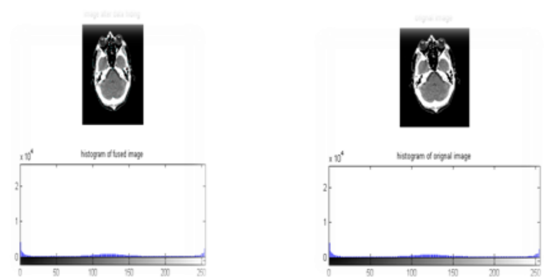
The mathematical representation of PSNR and MSE is given below:

$$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right)$$

$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}\|f(i,j) - g(i,j)\|^2$$

where m and n are the rows and columns of the image .f(i,j) is the original image and g(i, j) is the fused image.

Fig 3 shows original ctscan image and watermarked image



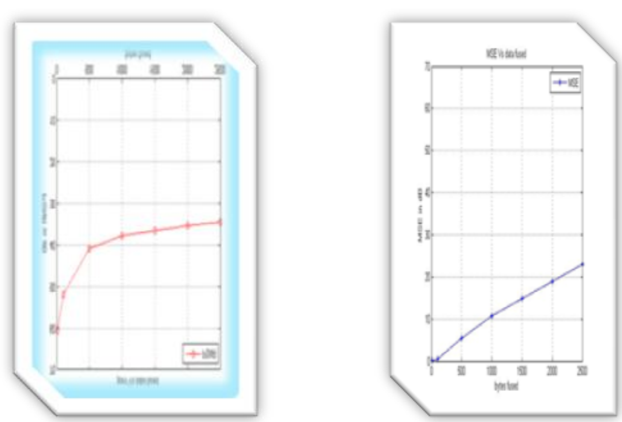
(a) (b)

Fig 3: (a) Original ctscan image and (b) Watermarked image

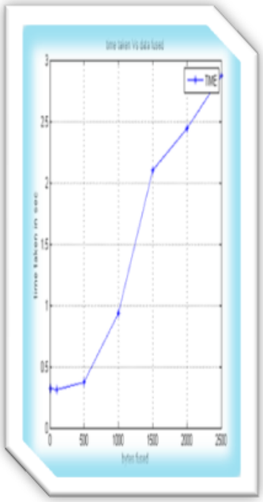
Table 1 below shows the variation of PSNR ,MSE and time taken to hide the patient data based on data capacity in bytes for ctscan image.then,fig 4 shows the graphical representation for variation of PSNR ,MSE and time taken vs data capacity.High value of PSNR shows the less image degradation .

Table 1: Variation of PSNR and MSE based data capacity (ctscan image**)**

Data fused(bytes)	PSNR(dB)	MSE(dB)	Time taken(sec)
10	60.5629	0.0576	0.3213
100	51.8303	0.4300	0.3099
500	40.7794	5.4769	0.3696
1000	37.8313	10.7981	0.9373
1500	36.4688	14.7776	2.1020
2000	35.3967	18.9150	2.4407
2500	34.5596	22.9361	2.8748



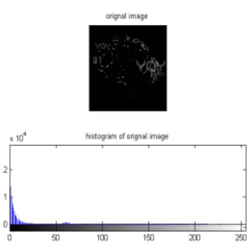
(a) (b)



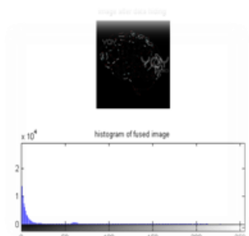
(c)

Fig 4: Graphs for ctscan images (a) PSNR vs data fused ,(b)MSE vs data fused and (c) bytes fused vs time taken to hide

Fig 5 shows the original brain image and watermarked image after hiding patient and table 2 shows the variation of PSNR, MSE and time taken to hide the patient data based data capacity.



(a)

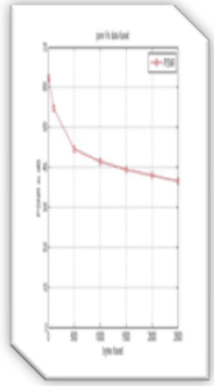


(b)

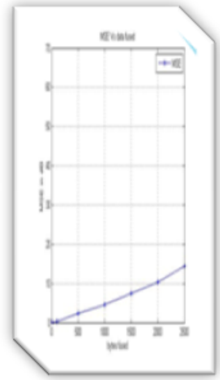
Fig 5: Brain images (a) original brain image, (b) fused brain image

Table 2: Variation of PSNR, MSE and time taken to hide data based on data capacity (Brain image)

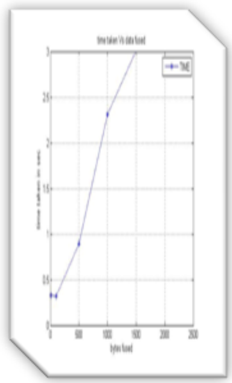
Data fused(bytes)	PSNR(dB)	MSE(dB)	Time taken(sec)
10	62.3068	0.0385	0.3316
100	54.7907	0.2175	0.3187
500	44.5106	2.3197	0.8918
1000	41.4401	4.7040	2.3093
1500	39.4699	7.4044	3.0069
2000	38.0103	10.3622	3.2750
2500	36.5773	14.4129	3.5736



(a)



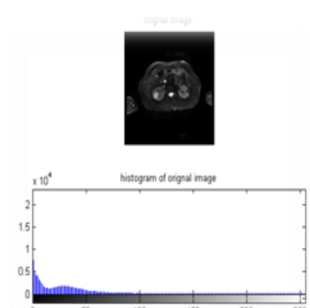
(b)



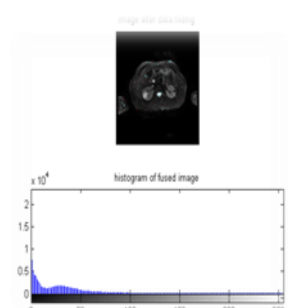
(c)

Fig 6: Graphs (a) PSNR vs data fused, (b) MSE vs data fused and (c) bytes fused vs time taken

Fig 7 shows the original liver image and fused image with size 512*512 and table 3 shows the below shows the variation of PSNR ,MSE and time taken to hide the patient data based on data capacity in bytes for ctscan image.



(a)

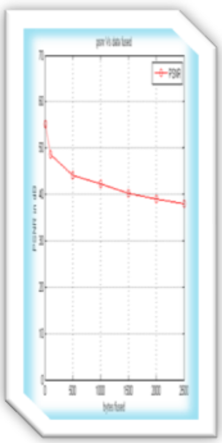


(b)

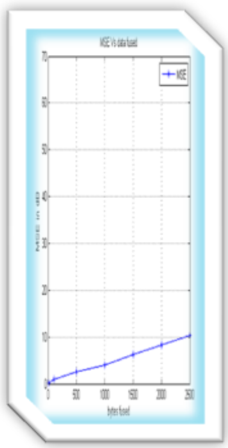
Fig 7 : liver images (a) original image , (b) fused liver image

Table 3: Variation of PSNR, MSE and time taken to hide data based on data capacity (liver image)

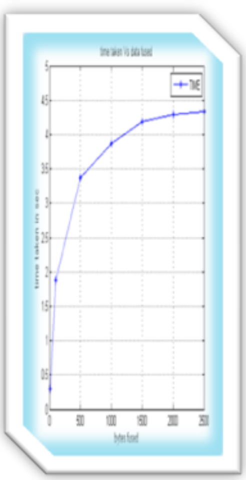
Data fused(bytes)	PSNR(dB)	MSE(dB)	Time taken(sec)
10	55.2368	0.1962	0.2988
100	48.6952	0.8850	1.8745
500	44.1132	2.5419	3.3767
1000	42.2025	3.9467	3.8616
1500	40.2472	6.1910	4.1885
2000	39.0200	8.2125	4.2970
2500	37.9861	10.4200	4.3339



(a)



(b)



(c)

Fig 8: Graphical representation (a) PSNR VS data fused , (b) MSE VS data fused and (c) time taken VS data fused

IV. CONCLUSION

Digital medical imaging technologies have become increasingly important in medical practice and health care for providing assistant tools for diagnosis: treatment, and surgery. Due to the volume of medical images is huge and has grown rapidly. Especially on CT (computer tomography) and MRI (magnetic resonance imaging), the compression technique must be applied. Our proposed work is very effective for hiding the large amount of patient data. As it finds the non area of interest in the medical image. Then, patient data is embedded onto noisy pixels. This algorithm is an effective algorithm providing good results with accuracy because it selects the region of non area of interest which is unusable part of the image. Hence there is not any problem at the time diagonosis of the image .Image does not lose its originality also. Results have shown the accuracy of non area of interest detection using fuzzy logic based matrix scanning algorithm over other algorithms. The fuzzy logic based algorithm has been successful in obtaining the noisy pixels that are present in an image after the its implementation and execution with various sets of images. Sample outputs have been shown to make the readers understand the accuracy of the algorithm. Thus developed algorithm exhibits tremendous scope of application in various areas of digital image processing.

IV. FUTURE SCOPE

1. In future, we will extend our proposed algorithm in order to obtain the less degradation and results with more accuracy .
2. We can use 5*5 and 7*7 Scanning mask window in future for predicting the noisy pixels in the medical image.
3. In our proposed work, the image is first to be converted into gray image. This limitation can be eliminated and algorithm can be applied directly to color images, and the detection would then become significantly more complex.

REFERENCES

[1] Hyung-Kyo Lee, Hee-Jung Kim, Ki-Ryong Kwon, Jong-Keuk Lee, "Digital Watermarking of Medical Image Using R01 Information" 2005 IEEE.

[2] Rodríguez-Colín Raúl, Feregrino-Uribe Claudia, Trinidad-Blas Gershom de J., " Data Hiding Scheme for Medical Images" 2007 IEEE .

[3] Liu Xin, Lv Xiaoqi, A Semi-fragile Digital Watermarking Algorithm Based on Integer Wavelet Matrix Norm Quantization For Medical ImagesWang Ying,"2009 IEEE.

[4] Henning M¨uller, Jayashree Kalpathy–Cramer," The ImageCLEFMedical Retrieval Task at ICPR 2010 Information Fusion,IEEE 2010.

[5] Xinde Sun and Shukui Bo, "A Blind Digital Watermarking for Color Medical Images Based on PCA" 2010 IEEE.

[6] Osamah M. Al-Qershi, Bee Ee Khoo," ROI–based Tamper Detection and Recovery for Medical Images Using Reversible Watermarking Technique",IEEE 2010.

[7] Hsiang-Cheh Huang A, Wai-Chi Fang B, Shin-Chang Chen "Privacy Protection and Authentication for Medical Images with Record-Based Watermarking" 2009 IEEE .

[8] P.Viswanathan, Dr.P.Venkata Krishna," Text fusion watermarking in Medical image with Semi-reversible for Secure transfer and Authentication", IEEE 2009.

[9] Nisar Ahmed Memon , Syed Asif Mehmood Gilani ," Adaptive Data Hiding Scheme for Medical Images using Integer Wavelet Transform ,IEEE 2009.

[10] Khurram Khurshid1, Claudie Faure2, Nicole Vincent," Fusion of Word Spotting and Spatial Information for Figure Caption Retrieval in Historical Document Images", IEEE 2009.

[11] P.Viswanathan, Dr.P.Venkata Krishna," Text fusion watermarking in Medical image with Semi-reversible for Secure transfer and Authentication", IEEE 2009.

[12] M.C. D´ıaz-Galiano, M.T. Mart´ın-Valdivia,A. Montejo-R´aez, L.A. Ure˜na-L´opez,"improving Performance of Medical Images Retrieval by Combining Textual and Visual Information"IEEE 2008.

[13] Amrinder Singh Brar, Mandeep Kaur, "Reversible Watermarking Techniques for Medical Images with ROI-Temper Detection and Recovery - A Survey" IJETAE 2010.

[14] Caroline Lacoste, Joo-Hwee Lim, Jean-Pierre Chevallet, and Diem Thi Hoang Le "Medical-Image Retrieval Based on Knowledge- Assisted Text and Image Indexing" 2007 IEEE .

[15] Joo-Hwee Lim, Caroline Lacoste, Jean-Pierre Chevallet, and Diem Thi," Hoang Le,"knowledge assisted medical image retrieval" IEEE 2007.

[16] Pedro R. Kalva, Fabricio Enembreck and Alessandro L. Koerich,"web Image Classification Based on the Fusion of Image and Text Classifiers"IEEE 2007.

[17] S.Poonkuntran, R.S.Rajesh, "A Messy Watermarking for Medical Image Authentication" 2011 IEEE.

AUTHORS

First Author – Pinki Jain, Doon Valley Institute of Engineering and Technology, Kurukshetra University, Karnal
Email id – pinki.jain1988@gmail.com

Second Author – Anu Aggarwal, Faculty of Doon Valley Institute of Engineering and Technology, Kurukshetra University, Karnal.
Email id - anu.cse.diet@gmail.com