International Journal of Pure and Applied Mathematics
Volume 118 No. 20 2018, 1975-1982
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

A FUZZY LOGIC BASED BIT PLANE COMPLEXITY SEGMENTATION STEGANOGRAPHY TO SECURE ELECTRONIC HEALTH RECORDS

G. Santhi[1], B. Adithya[2]
[1] Assistant Professor, [2] PG Student
[1,2]Department of Information Technology,
Pondicherry Engineering College, Pondicherry-605014, India
[1]shanthikarthikeyan@pec.edu, [2]adithya27.07@gmail.com

Abstract: Electronic Health Records or E-Health Records (EHR) is a paperless management promises to speed up typical bureaucracy of healthcare. A typical E-Healthcare system consists of many components and subsystem, such as appointment, routine clinical notes, picture archiving, lab and radiology orders, etc., are vulnerable to security threats. Cryptology and steganography are generally used to ensure medical data security. For this reason, a Fuzzy Logic based Bit Plane Complexity Segmentation (FL-BPCS) steganography is proposed to secure patient health records. In this, Electroencephalogram (EEG),Electrocardiography (ECG), Electromyography (EMG) time series, doctor's prescription and patient history are selected as hidden data and Magnetic Resonance (MR) image and Computed Tomography (CT) scan images in Digital Imaging and Communications in Medicine (DICOM) file format are used as cover image. The hidden data is compressed using Huffman lossless compression techniques and it is encrypted with AES to prevent from malicious attacks. The FL-BPCS algorithm is used to embed the encrypted data and stego image is created.

Keywords: E-Health Records, Steganography, FL-BPCS,AES, Stego image

1. Introduction

The privacy and security of Electronic Health Records information falls into two categories. First, inappropriare releases from authorized users who intentionally or unintentionally access or disseminare information violation of EHR system policies. Second, open disclosure of patient health information to parties against to the interests of specific individual patient or invadere a patient's privacy. These falls arises from the flow of data across the EHR system amongst and between providers, payers and secondary users with or without the patient's knowledge.

Technical obstaculum of the intrudere includes the use of firewalls to isolate internal networks with strong encryption based authentication and authorization. There is no known obstaculum for external networks like Denial-of-Service. Nevertheless, technical countermeasures cannot be cures all security threats. Obstaculum such as encryption, steganography and authentication are the only effective ways to counter EHR security threats against Internet interface.

Steganography is used for covert communication [2]. The embedding algorithm will convert the cover medium into stego medium by embedding secret data into it. The inverse process of embedding is done to extract the secret data. Imperceptibility, security, capacity, robustness, embedding complexity are the Steganography factors that has to be considered. Image steganography is developed according to its use in medical fields to communicate between patient as well as communication between doctor's and laboratory people's to hide secret messages. Steganography is to avoid drawing attention to the transmission of hidden medical information. If suspicion is raised, steganography and cryptography is planned to achieve the security of secret medical data's. Both are complementary to each other and provide better security, confidentiality and authenticity.

Nambakhsh et al. proposed a multi-resolution wavelet decomposition method to embed ECG and demographic text in sequential series pattern. This method is robust aganist the attacks but the capacity of this method is very low and it has very highly computational complexity [7][10].

Acharya et al proposed a Least Significant Bit (LSB) technique using Discrete Cosine Transform (DCT) of frequency domain, equivalent binary number are interleaved by the LSB of high frequency DCT coefficient. This method used the channel coding technique to decrease the effect of noise but there is an visible difference in histogram of cover and stego image [6].

Eiji Kawaguchi et al. proposed a Bit Plane Complexity Segmentation (BPCS) to increase the embedding capacity and also to overcome the short comings of traditional steganography techniques such as Least Significant Bit (LSB), Transform embedding, Perceptual masking techniques but embedded secret data can be retrieved by using Difference Image Histogram (DIH) [9].

Karakis et al. proposed a Fuzzy Logic based Least Significant Bit (FL-LSB) to reduce the insecurity of LSB planes. By using fuzzy the LSB planes were chosen to embed the secret data's but it has low embedding capacity and the stego image is invalid [12].

However, the imperceptibility is maintained in both spatial and transform domain. There is no accurate process is described to achieve robustness with high embedding capacity. Fuzzy logic is used to provide a feasible tool to solve the factors of steganography. Fuzzy logic exploits the pervasive imprecision information. So, adopting fuzzy logic to solve this problem is an appropriate choice. This paper proposes a Fuzzy Logic based Bit Plane Complexity Segmentation to secure the Electronic Health Records. It considers some parameters such as Peak Signal to Noise Ratio, Mean square Error, Structural Similarity Measure. The fuzzy logic choose the Most Significant Bit (MSB) and LSB bits to embed the secret data without affecting the DICOM cover image. The fuzzy logic is incorporated to avoid the distortion in the medical image because the image is has most sensitive data. It may change the Patients rights as well as Patient treatments.

This technique can be offered by using the following sections. In section 2, the brief description of BPCS, cryptology as well as lossless compression technique is presented, in section 3, the work methodology is presented, in section 4, results are presented, in section 5, the discussion and analysis are done, and finally the work is concluded.

2. Bit Plane Complexity Segmentation to Embed and Extract

This steganography method makes use of the human vision. The cover image is divided into informative region and noise-like region and the secret data is hidden in noise blocks of vessel image without degrading image quality. The data is hidden in both Most Significant Bit (MSB) as well as LSB planes. The important aspect of BPCS- steganography is high embedding capacity when compared to other traditional methodology. In LSB technique, data is hidden in last four bits. The multi-valued image (P) consisting of m-bit pixels can be decomposed into set of n binary pictures. Eg: P is an m-bit gray image say m=8. Therefore P= [P7, P6, P5, P4, P3, P2, P1, P0] where P7 is the MSB bit plane and P0 is the LSB bit plane [9]. Each bit plane can be segmented into informative and noise region. An informative region consists of simple pattern while noise-like region consists of complex pattern. In BPCS, each noise-looking region is replaced with another noise-looking pattern without changing the overall image quality [9].

3. Design of Fuzzy Logic Based BPCS

The proposed system aims to use EEG time series, doctor's comment, patient's information as secret data and MR images are used as cover or vessel image. The secret data's were compressed using Huffman lossless compression technique [1]. Second, the compressed data is taken as the input for encryption stage. To encrypt the compressed data the AES symmetric encryption technique [4] is used. Third, the encrypted data has to be embedded in cover image. To embed the encrypted data the FL-BPCS algorithm is used. Now, the stego image is created. The secret data is extracted from the stego image by using FL-BPCS algorithm. The extracted secret data which is in the encrypted form should be decrypted by using same AES symmetric algorithm. The decrypted data must be decompressed using same Huffman lossless compression technique. Now, the secret data is extracted as same as original embedded data. As mentioned in the above block diagram Fig.1, the data hiding and the data extracting will be done in two phases. (1) Embedding the secret data in the vessel image. (2) Extracting the secret data from stego image.

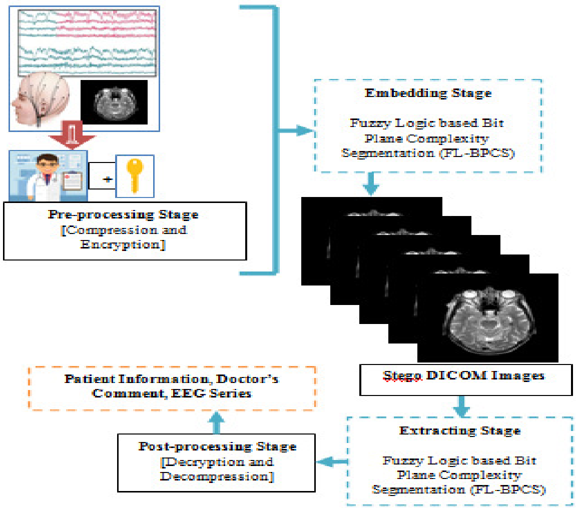

Embedding Stage
Fuzzy Logic based Bit Plane Complexity Segmentation (FL-BPCS)
+
Pre-processing Stage [Compression and Encryption]
Stego DICOM Images
Patient Information, Doctor's Comment, EEG Series
Post-processing Stage [Decryption and Decompression]
Extracting Stage
Fuzzy Logic based Bit Plane Complexity Segmentation (FL-BPCS)

Figure 1. System architecture of FL-BPCS

A. Embedding Phase

Least Significant Bit (LSB) and Most Significant Bit (MSB) embedding is a simple and fast strategy in steganography. It has high imperceptibility and embedding capability. Hence, this study proposes new methods to modify LSB and MSB embedding using medical data. The analyses consist of two stages: embedding and extracting, respectively. Initially, the patient's information is obtained from DICOM series of epilepsy patient. The EEG data is segmented according to the size of these DICOM images.

An image is sampled by pixels. In the gray-scale image, pixels have gray level intensities. In color images, pixels are also represented by three component intensities, being red (R), green (G), and blue(B). A similarity measure is the similarity degree between two groups or between two objects. Demirci [8] proposed a similarity-based method for edge detection. Furthermore, Pixel-Value Differencing (PVD) or Adjacent Pixel Difference (APD) methods determine embedding pixels in histogram-based steganography. Data Flow of overall proposed system process is shown in Fig. 1. These methods have high embedding capacity and PSNR values.

The main idea of this method is to generate a new image whose pixels have double values at the interval [0 1]. This similar image of cover image is used to determine pixels for the embedding message. In this method, if the values of similar pixels are higher than the determined threshold by trial and error, they are selected to hide the message [9]. The neighboring pixels of the image (P1,P2,...P9) using the 3x3 window have three color component (R, G, B).

The gray level differences of color components are calculated the neighboring pixels of the stego image. The color distance of pixels are calculated by the Euclidean norm [12]. The similarity values of pixels are founded. Similarly, the coordinates of the pixels are determined between the similarity values of pixels and threshold values. The hidden message is extracted using the coordinates of the stego image's pixels.

B. Fuzzy Logic based Bit Plane Complexity Segmentation

The system adjust the pixel value of an medical image by comparing the vessel image value and the extracted image value. It use 5-level fuzzy logic system.

Step 1: Linguistic variables and terms of fuzzification
Linguistic variables are input and output variables in the form of simple words or sentence. For vessel image, Gray level differences (Red, Green, Blue) of pixels are Purely Differ (PD), Moderately Differ (MD), Exactly Similar (ES), etc., are linguistic terms. Every member of this set is a linguistic term and it can cover some portion of overall pixel values. Three linguistic terms are only used, because we were using both MSB as well as LSB bits in MR image. Already, the MR image is in high quality. So, any variation in the pixel value is noticed at this level.

Step 2: Membership functions
A membership function (MF) is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. Three triangular fuzzy sets are used to fuzzify the gray level differences. But according to my research and knowledge in this field, these shapes are simple and more flexible. So, choosing membership requires expert knowledge. The membership functions of gray level differences is shown in Fig.2.

Membership Function

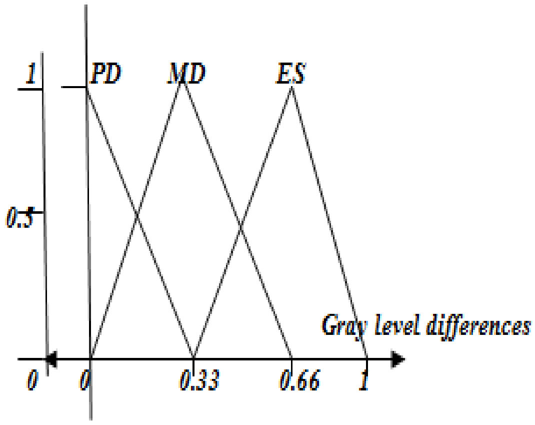


Figure 2. Fuzzy membership values of gray levels of red, green, and blue components for neighbor pixels

Step 3: Knowledge base rules
Build a set of rules into the knowledge base in the form of IF-THEN-ELSE structures.
Rule 1: If RED is Low and GREEN is Low and BLUE is Low, Then P_1 and P_2 are Exactly Similar.
Rule 2: If RED is High and GREEN is High and BLUE is High, Then P_1 and P_2 are Purely Differ.
Rule 3: If RED is High (H) and GREEN is Low (L) and BLUE is Medium (M), Then P_1 and P_2 are Moderately Differ. Since each input variable has three linguistic states, the total number of possible inference rules is 3*3=27. Table.1 shows some fuzzy rule base.

Table 1. Fuzzy rules of the BPCS

RED	GREEN	BLUE	μ_A
L	L	L	ES
L	L	M	ES
L	M	H	MD
M	H	L	MD
M	H	M	ES
M	H	H	PD
H	H	L	ES
H	H	M	PD
H	H	H	PD

Step 4: Obtaining fuzzy value

Fuzzy set operations perform evaluation of rules. The operations used for OR and AND are Max and Min respectively. Combine all results of evaluation to form a final result. This result is a fuzzy value.

Step 5: Defuzzification

Defuzzification is the process of converting a fuzzified output into a single crisp value with respect to a fuzzy set. Defuzzification is then performed according to membership function for output variable (**μ_A**). This method provides a crisp value based on the center of gravity of the fuzzy set. Eq.1, the area and the center of gravity or Eq.2, centroid of each sub-area is calculated and then the summation of all these sub-areas is taken to find the defuzzified value for a discrete fuzzy set.

$$A = \sum_{j=1}^{n} A_j \, \mu_{prem}^{j} (L) \div \sum_{j=1}^{n} \mu_{prem}^{j} (L) \qquad (1)$$

$$\mu_{prem}^{j} (L) = \mu_R^j \, (RED) \, \mu_G^j \, (GREEN) \, \mu_B^j \, (BLUE) \qquad (2)$$

C. Extracting Phase

Huffman Compression technique is used to increase message capacity. To increase security, the compression message is encrypted by the Rijndael symmetric encryption algorithm using a 128-bit key. Secondly, the proposed methods, which are Fuzzy-Logic-based Bit Plane Complexity Segmentation (FL-BPCS), select MSBs and LSBs of image pixels with using the differences in gray levels of the pixels. Finally, the selected MSBs and LSBs of the pixels are altered with the message bits in stego images. These processes are simultaneously run with all DICOM series to decrease computational time.

The extracting message stage requires stego-DICOM images is shown in Fig. 3a, 3b and a stego-key, which is the authentication key for decryption. Firstly, the proposed methods give the pixels coordinates, which have an embedded message. These pixels are used to gather the message. Secondly, the obtained message is decrypted and decompressed. Finally, the patient's information, segmented EEG, and the doctor's comments are displayed in a GUI (Graphical User Interface) screen. All hidden EEG data can be also gathered from the DICOM series. The comparison results of the proposed algorithm are evaluated by Histogram is shown in Fig. 4a, 4b, PSNR (peak signal-to-noise ratio), MSE (mean square of error), SSIM (structural similarity measure), between the cover, and the stego-DICOM series.

4. Results Discussion and Analysis

The proposed system parameters are used to evaluate the performance of the data hiding techniques.

Peak Signal to Noise Ratio (PSNR): The PSNR is generally used to measure the quality of stego image in decibels (dB). Eq.3, gives the expression for PSNR in which I_{Cmax} is the maximum pixel value of the cover image and MSE is the mean square error:

$$PSNR = 10 \log_{10} \left(\frac{I_{Cmax}^2}{MSE} \right) dB \qquad (3)$$

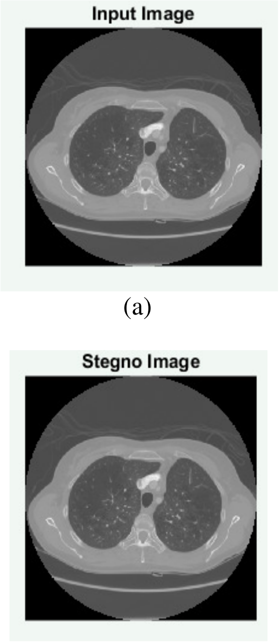
Where,

$$MSE = \frac{1}{MN} \sum_{X=1}^{M} \sum_{Y=1}^{N} \left(I_{Sxy} - I_{Cxy} \right)^2 \qquad (4)$$

In Eq. 4, x and y are the image coordinates, M and N are the dimensions of the image, I_{Sxy} is the generated stego-image and I_{Cxy} is the cover image [11] [13].

Structural similarity (SSIM) index: The SSIM is a method for finding the similarity between cover image and the stego image. It is a perception-based model that considers image degradation as perceived change in structural information. The SSIM measure between two images I_c and I_s is represented in Eq. 5, where, μ_{I_c} is the average of I_c, μ_{I_s} is the average of I_s, $\sigma_{I_c}^2$ is the variance of I_c, $\sigma_{I_s}^2$ is the variance of I_s, σ_{I_c, I_s} is the covariance between I_c and I_s and k₁, k₂ are two the variables used to stabilize the division with weak denominator [13].

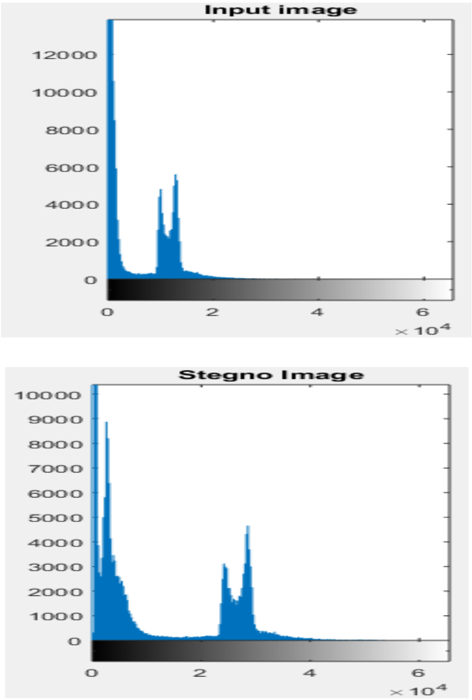
$$SSIM \langle I_c, I_s \rangle = \left(\frac{\langle 2\mu_{I_c} \mu_{I_s} + K_1 \rangle \langle 2\sigma_{I_c, I_s} + k_2 \rangle}{\langle \mu_{I_c}^2 + \mu_{I_s}^2 + k_1 \rangle \langle \sigma_{I_c}^2 + \sigma_{I_s}^2 + k_2 \rangle} \right) \qquad (5)$$



(a)

(b)

Figure 3. (a)Cover image (b)Stego image



(a) (b)

Figure 4. Histogram of (a)Cover image (b)Stego image

The graph is created based on the *embedding capacity* is shown in Fig. 5. For embedding payload, we use an embedding rate, to represent the percentage of the embedded secret bits in the whole pixels of the cover image [11]. In Fuzzy Logic based Least Significant Bit the embedding capacity is mentioned as low, because to embed the secret data it occupies only the LSB positions. In Bit Plane Complexity Segmentation the embedding capacity is mentioned as low, when compared to the Fuzzy Logic based Least Significant Bit but normally the embedding capacity is high when compared to the primitive LSB because to embed the secret data it occupies both MSB as well as LSB. In Fuzzy Logic based Bit Plane Complexity Segmentation the embedding capacity is high because to embed the secret data the red, green, blue channel were used with fuzzy.

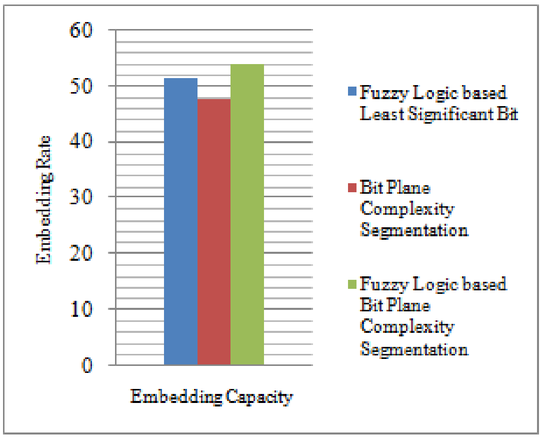


Figure 5. Embedding capacity

The graph is created based on the performance evaluation parameters [*PSNR, MSE, SSIM*] is shown in Fig. 6. Peak signal-to-noise ratio [3], often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR is most easily defined via the mean squared error (MSE). Two of the error metrics used to compare the various image compression techniques are the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR). The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error [14]. Structural Similarity Index (SSIM) is a perceptual metric that quantifies image quality degradation caused by processing such as data compression or by losses in data transmission. It is a full reference metric that requires two images from the same image capture--- a reference image and a processed image [5].

In Fuzzy Logic based Least Significant Bit the PSNR is achieved high, because to embed the secret data it occupies only the LSB positions and MSE is achieved low when compared to Bit Plane Complexity Segmentation and SSIM is achieved high. In Bit Plane Complexity Segmentation the PSNR is achieved as low because to embed the secret data it occupies both MSB as well as LSB and MSE is achieved high and SSIM is low because MSB is more sensitive. While changing that pixel values with another pixel values must be matched. In Fuzzy Logic based Bit Plane Complexity Segmentation the PSNR is low because to embed the secret data MSB as well as LSB is used. MSE is low and SSIM is high because it uses red, green, blue channels with fuzzy rules.

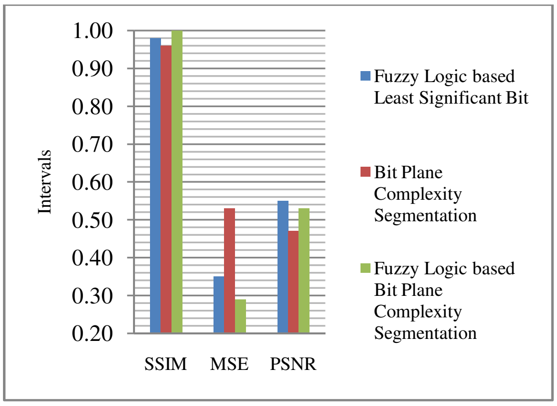


Figure 6. Performance Parameters

The graph is created based on the *steganalysis* is shown in Fig. 7. Visual attacks involve observing the unusual patterns and noisy blurred regions in some places of the stego image. A statistical method called RS steganalysis for detection of LSB embedding uses dual statistics derived from spatial correlation of an image. Histogram based steganalysis techniques detect the existence of secret data from smoothness of the stego image histogram. Similarly, a targeted active steganalysis technique is implemented for HS embedding using the change in the characteristics of histogram during data embedding [13]. In Fuzzy Logic based Least Significant Bit the visual attacks, RS statistical attack, Sample Pair Analysis is achieved high, because to embed the secret data it occupies only the LSB positions but Difference Image Histogram is achieved low. In Bit Plane Complexity Segmentation the visual attacks, RS statistical attack, Sample Pair Analysis is achieved low, because to embed the secret data it occupies both MSB as well as LSB but Difference Image Histogram is achieved high. In Fuzzy Logic based Bit Plane Complexity Segmentation the visual attacks, RS statistical attack, Sample Pair Analysis is achieved low, because to embed the secret data the red, green, blue channel were used with fuzzy rules but Difference Image Histogram is achieved high.

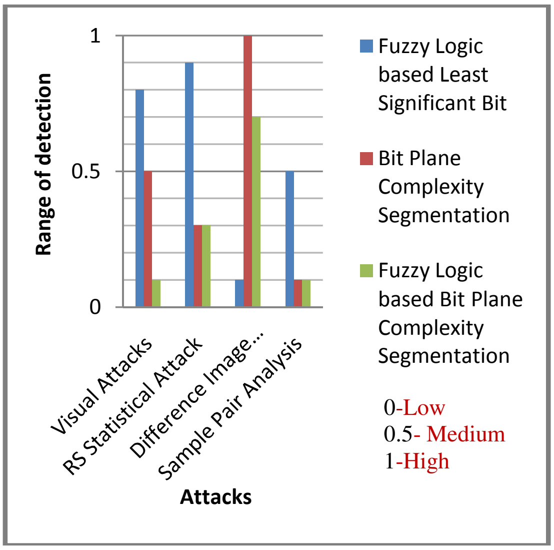


Figure 7. Detection of Embedded Data's

5. Conclusion

In medical information system, medical data is easily captured when being storing, receiving or transmission through computer network and Internet. Cryptology and steganography are generally used to ensure medical data security. For this reason, this study proposes new algorithm, Fuzzy Logic-based Bit Plane Complexity Segmentation (FL-BPCS) to secure medical data. The embedding messages are composed of EEG signals, doctor's comment, and patient information in file header of DICOM images. The messages are secured by using Huffman lossless compression methods and Rijndael symmetric algorithm with 128 bit-key to prevent the attacks. The capacity of proposed algorithm is higher than the result of similar studies in literature. According to the obtained result, the proposed method ensures the confidentiality of the patient's information. The FL-BPCS method hides EEG signals, patient's information and doctor's comment in the pixels of MR images. It also reduces data repository and transmission capacity of the patients' multiple medical data.

References

[1] Huffman, D. , "A Method for the Construction of Minimum-Redundancy Codes", Proceedings of the IRE. 40 (9):1098–1101, (1952).doi:10.1109/JRPROC. 1952.273898

[2] Simmons, G.J., “The Prisoners' Problem and the Subliminal Channel”, In: Advances in Cryptography, Chaum, D. (Ed.). Springer, New York, USA., ISBN-13: 9781468447323, pp: 51-67, (1984).

[3] Welstead, Stephen T,” Fractal and wavelet image compression techniques”, SPIE Publication, pp. 155–156, (1999). ISBN 978-0-8194-3503-3

[4] Daemen, Joan; Rijmen, Vincent, "AES Proposal: Rijndael", National Institute of Standards and Technology, p. 1. March 9, (2003).

[5] Wang, Zhou; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P., "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing**,** 13 (4): 600–612, (2004-04-01). doi:10.1109/TIP.2003.819861

[6] Acharya U. R., Niranjan, U.C., Iyengar S.S, Kannathal, N., Min L, “Simultaneous storage of patient information with medical images in the frequency domain”, Computer Methods and Programs in Biomedicine, Volume 76, Issue 1, 76 (1), 13–19, (2004).

[7] Nambakhsh, M.S., Ahmadian, A., Ghavami, M., Dilmaghani, R.S., Karimi-Fard. S,”A Novel Blind Watermarking of ECG Signals on Medical Images Using EZW Algorithm”, Proceedings of the 28th IEEEEMBSAnnual International Conference New York City, USA, 3274-3277, (2006).

[8] Recep Demirci, “Similarity relation matrix-based color edge detection”, AEU-International Journal of Electronics and Communications, vol 61, issue 7, pages: 469-477, (2007).

[9] Souvik Bhattacharyya, Aparajita Khan,Aunkita Nandi,Aveek Dasmalakar,Somdip Roy and Gautam Sanyal, “Pixel Mapping Method (PMM) Based Bit Plane Complexity Segmentation (BPCS) Steganography”, Information and Communication Technologies, pages: 36-41, (2011).

[10] Nambakhsh, M.S., Ahmadian, A., Zaidi, H, “A contextual based double watermarking of PET images by patient ID and ECG signal”, Computer Methods and Programs in Biomedicine, 104:3, 418-425, (2011).

[11] Yanping Zhang, Juan Jiang, Yongliang Zha, Heng Zhang, Shu Zhao, “Research on Embedding Capacity and Efficiency of Information Hiding Based on Digital Images”, SciRes.., International Journal of Intelligence Science, 77-85, (2013).

[12] R. Karakış, İ. Güler, İ. Çapraz and E. Bilir, “A Novel Fuzzy Logic-Based Image Steganography Method To Ensure Medical Data Security”, Computers in Biology and Medicine, vol 67, pages: 172-183, (2015).

[13] Ahmad Shaik, V. Thanikaiselvan and Rengarajan Amitharajan,. “Data Security Through Data Hiding in Images: A Review”, Journal of Artificial Intelligence, 10: 1-21, (2017).

[14] From Wikipedia, the free encyclopedia, “Peak signal-to-noise ratio”, https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio.

[15] T. Padmapriya and V. Saminadan, “Improving Throughput for Downlink Multi user MIMO-LTE Advanced Networks using SINR approximation and Hierarchical CSI feedback”, International Journal of Mobile Design Network and Innovation- Inderscience Publisher, ISSN : 1744-2850 vol. 6, no.1, pp. 14-23, May 20 15.

[16] S.V.Manikanthan and T.Padmapriya “Recent Trends In M2m Communications In 4g Networks And Evolution Towards 5g”, International Journal of Pure and Applied Mathematics, ISSN NO: 1314-3395, Vol-115, Issue -8, Sep 2017.