

Research article

Dual Watermarking for Protection of Medical Images based on Watermarking of Frequency Domain and Genetic Programming

Abdul Joseph Fofanah

Milton Margai College of Education & Technology
Department of Mathematics and Computer Science

E-mail: abduljoseph.fofanah@gmail.com

Ibrahim Kalokoh

University of Makeni
Department of Computer Science

E-mail: ikalokoh@gmail.com



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

ABSTRACT

In this paper, we presents two watermarking schemes namely: Scheme I present an hybrid of discrete cosine transform (DCT) and discrete wavelet transform (DWT) watermarking technique while Scheme II presents a genetic programming (GP) based technique. Both of these schemes were experimented with 512x512 and 256x256 medical images and other standard images such as Lena, Barbara, Baboon, Flower, Brain etc. against state-of-the-art watermarking techniques. According to our results, shows better performance of the two main criteria use to check for any watermarking techniques (Imperceptivity and Robustness). The experimental results reveals that for imperceptivity, peak-signal-to-noise-ratio (PSNR) and structural similarity index (SSIM) shows better performance when compared against Arsalan's and Hurrah's approaches. Similarly, for robustness the parameters used are normalized correlation (NCC) and bit error rate (BER). According to results obtained, both schemes' parameters show high performance except for Scheme I BER with 0.0038. Finally, in terms of fragility both schemes provided better resistance to any signal processing attacks.

Keywords: Discrete Cosine Transform (DCT); Genetic Programming (GP); Discrete Wavelet Transform (DWT); evolutionary algorithms; watermarking techniques.

1. INTRODUCTION

The development of high-intensive computer networks and the acceptance of electronic management system in medical records system have made it possible for digital medical images to be shared across the globe essentially in

services such as teleconsultation¹[1], telemedicine, tele-diagnosis and teleradiology² [2-6]. Automatic diagnosis and understanding of a certain disease and minimizing the number of misdiagnosis has had an extensive social and economic impact. This essentially shows the need for robust and efficient patient information sharing between professional of different hospitals. In the management of medical images, the main precedence is to secure protection for the patient's documents against any act of altering by unauthorized persons. Entirely these technological developments currently presented a progressive change in different healthcare.

Facilities such as hospital management information system (HMIS), information management, medical imaging and specially health social networks [8],[15]. According to [22], Telemedicine is defined as the use of information and communications technologies (ICTs) in order to provide healthcare services when medical practitioners, patients and researchers are present in various geographical locations.

However, the transmission and sharing of electronic patient record (EPR) raises different security related challenges such as reliability, authenticity, security, integrity and confidentiality as discuss by researchers in this field such as [9], [17], and [22] respectively. digital cinema and content authentication are some of the major applications of digital watermarking ([1], [5], [7],[9] , [17], [22], [18],[23] and [22]).

Essentially, digital watermarks are also used to secure state driver licenses by providing machine readable and convert layer of security to combat against security challenges such as fraud, identity theft and counterfeiting. Equally medical image watermarking (MIW) has diverse advantages such as save storage space and band-width requirements, confidentiality of patient data [22]. In addition, according to [31] challenges of various security concerns reported in various survey asserted that medical image watermarking also helps in reducing medical identity thefts. Essentially, the major drawbacks of digital image watermarking schemes is that the embedded watermarked document should not affect the user perceptivity quality of the covered image and should be robust and devoid of different signal processing attacks.

According to [11], in 2018 report, it was found to have the prevalent healthcare data infringements and accounted for the following parameters: loss, theft, unauthorized accessing, impermissible disclosure or improper disposal of 100,000 or more healthcare records. The re-port indicated that in 2018 has seen 18 data infringements that has been open to attacks of 100,000 or more healthcare records attacks and three of those infringements attacks greater than 1 million healthcare records .The Department of Health and Human Services' Office of Civil Rights (OCR) has received reports of 365 data infringements of 500 or more EPRs.

2. RELATED WORKS

According to Kirchho 's principle, the security of an encryption and cryptographic technique it highly dependent on the secret key [30]. Therefore, since the cryptosystem is dependent only on the secret keys, it should be highly impossible to retrieve the information without the authentication of the exact secret keys. On the other hand, [16] presented the encryption timing analysis for an IoT based data security system and facts shows that encryption techniques used like AES, Kurupira and Trivium have a comparatively lesser computational overhead compared to the dual and multiple watermarking algorithm. The increased encryption time in this case could be attributed to use of double layer encryption (Arnold transform followed by dual watermarking encryption algorithm).

Furthermore, [4], [29] in genetic programming the based intelligent coefficient selection is performed in integer wavelet domain, to exploit the inter-dependencies of wavelet coefficient. For copyright protection [14], reported histogram-based techniques, in order to add robustness, while [2] proposed histogram shifting and clustering based watermarking technique that is not only robust but also reversible. [21] proposed a watermarking algorithm which is reversible and dependent on the prediction error-based expansion. In their attempted to solve the problem, the approach, bits are embedded based on the scale of variation in the neighboring eight-pixel values; hence there is no need for the location map or histogram shifting.

With cumulative increasing occurrence of cloud computing and storage, businesses and other institutions are rapidly increasing the trend of hosting large data storage and management into the cloud systems [25]. [32], proposed

¹ A general term use for any consultation between medical doctors and patients on a network or video link such as Facetime, intranet, internet, or skype

² The transmission of radiological patient images, such as x-rays, CTs, and MRIs, from one location to another for the purposes of sharing studies with other radiologists and physicians". Wikipedia

a system named SEISA with access control and se-cure k-means outsource, dynamically updating images is reinforced. In another research pro-posed by [27], the scheme can explore user relationship whilst preserving the privacy of the image. that uses the one-way hash encryption partly hash values to encrypt the image characteristics was another schemed proposed to enhance the protection and privacy issues in the cloud system [26]. This scheme created trade-o s among privacy preserving, quality retrieval, and complexity through adjusting the bit counts of encryption in the hash value. However, the scheme has some shortcomings in terms of imperceptivity and robustness of the embedded image.

According to [13] proposed a region based watermarking algorithm related to discrete wavelet transform and spread spectrum method. The medical cover image is disintegrated into two namely; ROI and NROI. The binary watermark is embedding into DCT transform of NROI portion of the cover using spread spectrum embedding technique. [19] presents a compression and encryption based data hiding technique using image moment theory for medical images. Encryption based medical images watermarking technique using LSB and DWT was also pro-posed by [12]. The medical image watermark is embedded in each block of cover image in the designated DWT sub-band using LSB. In healthcare field, the security of EPR data is primary concern to protect the confidential patient reports from the unauthorized access and unsolicited alterations. Therefore, [20] pro-posed DWT and DCT based multiple watermarking method using RSA and hamming error correcting codes. The image and EPR watermark is embedded simultaneously into the cover for the purpose of ownership identification and enhanced security of the medical information by identifying ROI and NROI of the cover image. The EPR and image watermarks are hidden into the NROI and ROI part of the cover medical image respectively. The proposed hybrid (DWT and DCT) watermarking method enhanced the NCC and PSNR performance as compared to DWT and DCT applied individually [20]. Additionally, RSA and MD5 are applied on EPR and image watermark respectively before embedding into the cover, which provides the extra level security of the water-marks.

3. PROPOSED WATERMARKING TECHNIQUE

In this paper, we present two watermarking schemes for combined DCT+DWT techniques and using an intelligent GP technique for medical image watermarking. Again to test an avail-able image authentication methods will be done and the suitable ones which have the ability to be used in the proposed watermarking technique (Scheme 1 and 2). Furthermore, the most reliable watermarking algorithm to use in the current state-of-the-art schemes will be identified. The main problem of the spatial do-main is that the methods are not sufficiently robust. Robustness is very important during implementation in the cloud environment to as-sure the resilience of the watermark. Comparatively, frequency domain methods provide more robustness that is more suitable to be used in the cloud environment. The frequency do-main methods: Singular Value Decomposition (SVD) and Distributed/Discrete Wavelet Trans-form (DWT) are more appropriate for robust image watermarking and also in copyright protection. The SVD technique enhances the robustness of the watermark against geometric and non-geometric attacks and can also be used for image watermarking. Another sufficient robust method for image watermarking, in digital image le formats is DWT. DWT is a method, which is used for frequency domain techniques. Consequently the concept of DCT, and DWT method will be explored as implemented in this paper work.

3.1 Scheme I embedding process

Firstly, we apply DWT to divide the cover host image into four non-overlapping multi-resolution sub-bands: LL1, HL1, LH1, and HH1. Next, we employ DWT again to sub-band HL1 to get four smaller sub-bands and select the HL2 sub-band as indicated in figure 1 . Or, employ DWT to sub-band HH1 to get another four smaller sub-bands and select the HH2 sub-band. Next step, is to divide the sub-band HL2 (or HH2) into 8 x 8 blocks. Henceforth we employ DCT to each block in the chosen sub-band using HL2 or HH2. Furthermore we re-formulate the grey-scale watermarked image into a vector of zeros and ones. The next stage is to generate two uncorrelated pseudorandom sequences. One sequence is used to embed the watermark bit 0 (i.e. PN0) and the other sequence is sued to embed the watermark bit 1 (i.e. PN1). Number of elements in each of the two pseudorandom sequences must be equal to the number of mid-band elements of the DCT-transformed DWT sub-bands. Next, we embed the two pseudorandom sequences, PN0 and PN1, with a gain factor, in the DCT transformed 8 x 8 blocks of the selected DWT sub-bands of the host image. It should be also noted that the gain factor determines the performance of PSNR, the higher the gain factor the lesser the PSNR and lesser gain factor the more robust the algorithm. Next, we apply inverse DCT (IDCT) to each block after its mid-band coefficients have been modified to embed the watermark bits as described in the previous step. Additionally, we apply the inverse DWT (IDWT) on the DWT transformed image, including the modified sub-band, to produce the watermarked host image.

3.2 Scheme I extraction process

After we have succeeded the embedding process, now we can employ DWT to modularize the watermarked image. To do this the LL1, HL1, LH1 and HH1 which makes up to the four non-overlapping multi-resolution. Next we employ DWT to HL1 to get four smaller sub-bands, and select the sub-band HL2. Next stage, we apply DWT to the HH1 sub-band to in order to achieve four smaller sub-bands, and select the HH2 sub-band, as shown in 2. Next, we divide the sub-ban HL2 (or HH2) into 4x4 blocks. Furthermore, to apply DCT to each block in the selected sub-band of HL2 or HH2 and then extract the mid-band coefficients of each DCT transformed block. Next we need to regenerate the two pseudorandom PN1 and PN0 sequences as described in watermarked embedding process Now, to compute the correlation between the mid-band coefficients and the two (PN0 and PN1) generated pseudorandom sequences. If the correlation with the PN0 was higher than the correlation with PN1, then the extracted watermark bit is considered 0, otherwise the extracted watermark is considered 1. Lastly, to re-construct the original image using the extracted watermark bit, and other parameters use to evaluate the difference between watermarked image and original image, we compute the similarity between the two images.

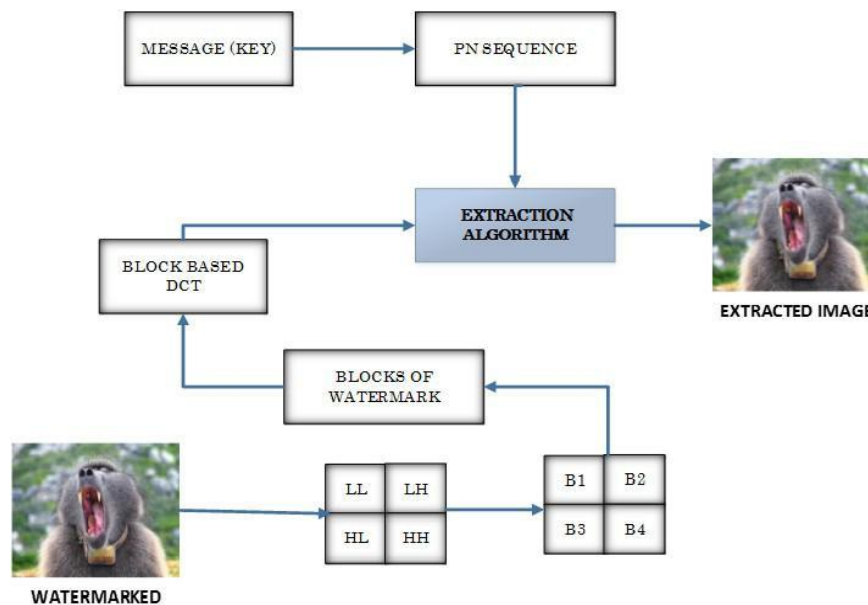


Figure 2: Schematic diagram of watermark extraction phase of the proposed DCT-DWT technique (Scheme I)

3.3 Description of Scheme II

GP Parameter Setting: In order to represent the candidate solution, we explore the GP terminal and function sets which are defined in table 1. The variables or parameters mentioned in table 1 use to generate expressions by choosing a suitable coefficients and exploring the neighborhood and each dependency. GP arithmetic operators (+, -, *, x), log, sin, cos, max, and min formed the GP functional set and to compute the structural complexity of the watermarked image.

3.4 Medical image watermark embedding and extraction phase using GP model:

In this section, we are exploring on embedding and extraction process and two main process involve. The value of the expression denoted as x is computed during the training phase. The value of x is used for the embedding of watermark during the training and testing phase. Next is to recover the message (key) by performing the extraction of the message which is reversible along with the embedded watermark.

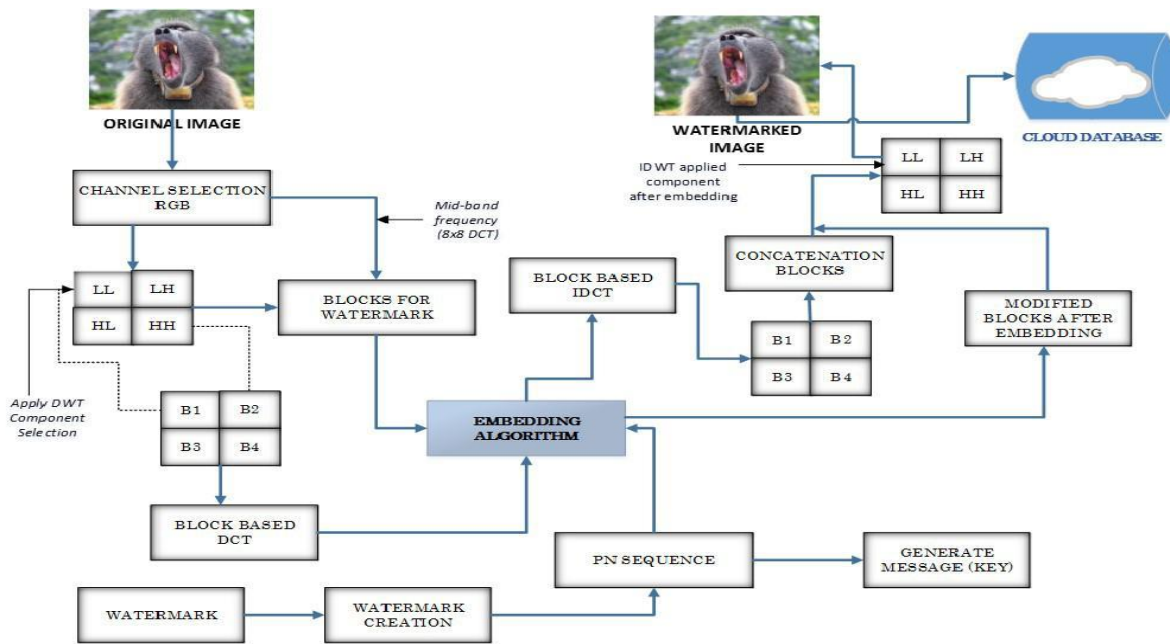


Figure 1: Schematic diagram of watermark embedding phase of the proposed DCT and DWT technique (Scheme I)

3.5 Scheme II embedding process

The program first read the original image and also ready the secret image before embedding. Then perform DWT transformation (i.e. LL, LH, HL, HH and LL1, LH1, HL1, HH1). Next determine the gain factor for embedding and set the DCT block size (block size = 4 and the gain factor used in this research is 50). Hence-forth defines the mid-band frequencies of an 8 x 8 DCT.

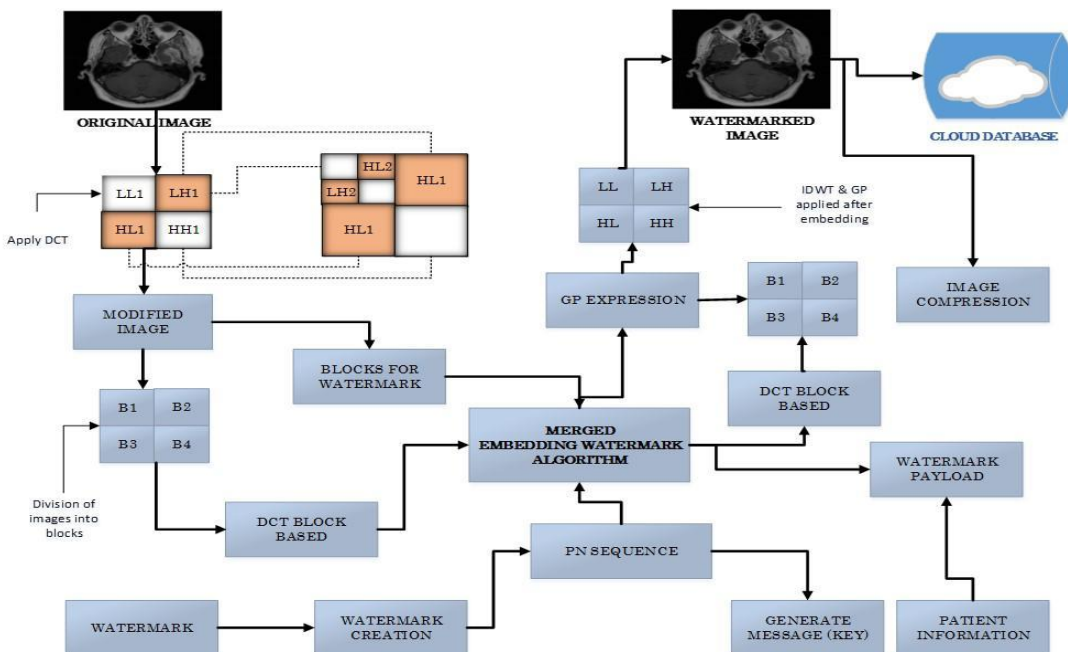


Figure 3: Schematic diagram of watermark embedding phase of the proposed GP technique (Scheme II)

Apply the original image HL1 and determine the size of cover image before embedding. Next, determine the maximum message size based on cover object (original image) through its height and width. Reshape the message to a vector of 256 by checking that the message is not too large for embedding. Apply padding message out to the maximum messages size with ones. Generate shell of watermarked image and read in key or message for PN generator and store it in the database, otherwise reset MATLAB's PN generator to state key and generate PN sequence. The next stage is to process the image in blocks by transforming block using DCT. If the message bit contains zero then embed PN sequence zero into mid-band components of the DCT block. Furthermore, apply transform block back into spatial domain and move on to the next block at and of row move to next row and finally display the watermarked image and compare it to the original image. The GP expression is used to compute the structural complexity of the watermarked image of the con-catenated DWT and GP expression (see figure 3 complexity of the algorithm).

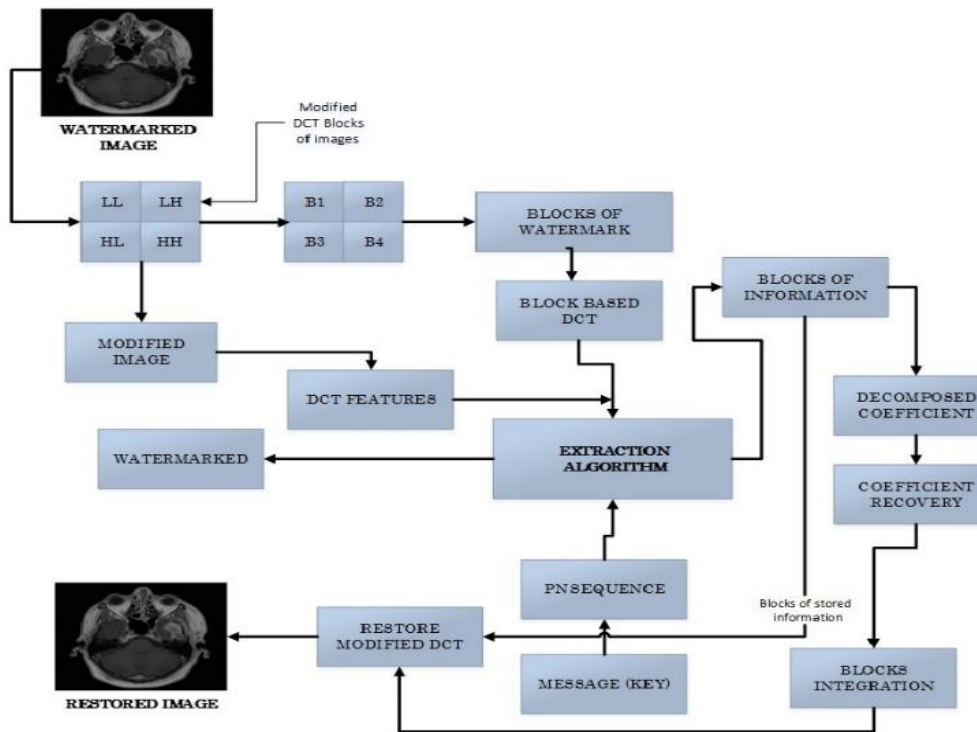


Figure 4: Schematic diagram of watermark extraction phase of proposed GP technique (Scheme II)

3.6 Scheme II extraction process

The extraction process can be achieved by setting the DCT block size as indicated in the embedding phase and defines the mid-band frequencies of an 8 x 8 DCT as shown in figure 4. Apply detection by reading the watermarked image of the same dimension and determine the maximum message size based on cover object, and block size and read in original watermark. Next needed to determine the size of original watermark and reset MATLAB's PN generator to state key. Display the generated PN sequence and process the image in block wise. Again in order to extract the middle band coefficients we need to transform block using DCT. Next stage is to calculate the correlation of the middle band sequence to PN sequence and move on to the next block in row wise sequence. If the correlation exceeds threshold, set bit to '0' otherwise '1' and reshape the embedded message and display the covered message. The block diagram of the extraction phase is shown in figure 4.

3.7 Parameters used for proposed schemes

In this section, we present a brief description of the parameters use to determine the quality and capacity of an image watermarked.

3.7.1 Normalized Cross Correlation: The normalized cross correlation or correlation coefficient (NCC) can be used to measure the compatibility between the original and extracted water-mark image. The minimum and maximum values of this matrix are 0 and 1 respectively. This is illustrated in equation (1).

$$\frac{\sum_{i=1}^M \sum_{j=1}^N O_w(i,j) * W_m(i,j)}{\sum_{i=1}^M \sum_{j=1}^N O_w(i,j)^2} \quad (1)$$

Where O_w is the original or cover image, W_m the watermarked image, i and j are the pixels of the images.

$$\frac{\sum_{i=1}^M \sum_{j=1}^N O_w(i,j) * W_m(i,j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N O_w(i,j)^2} * \sqrt{\sum_{i=1}^M \sum_{j=1}^N W_m(i,j)^2}} \quad (2)$$

Equation (2) is used to find the correlation coefficient of the image.

Table 1: Terminal set for GP Setting

Variable	Description
Sub_value	The mathematical value assigned to the different wavelet sub-bands used for embedding of watermark i.e. LH, HL, HH.
Sub_mean	The average value of coefficients in a specific sub-band.
Block_num	Block position within a sub-band
Block_max	The maximum coefficient value in a specific block.
(p, q)	p and q represent for the row and column indices of the coefficient in a specific block.

3.7.2 Bit Error Ratio: The bit error ratio (BER) is a useful evaluator when the watermark is a binary sequence. BER shows the probability of binary patterns that are decoded incorrectly. Therefore, the lower the value of BER , the better is the performance of the watermarking system. It is given by the equation (3).

$$BER = \frac{DB}{NB} \quad (3)$$

Where DB is the number of bit that is de-coded incorrectly and N B is the total number of original watermark bits. Since the BER basically identifies the average probability of incorrect bit identification. Thus a bit of 10^{-10} means that 1 bit out of every 10^{10} bits is , on average, read incorrectly. If the system is operating at 100Mb/s that is 10^8 pulses per second then receive 10^{10} pulses the time taken would be $\frac{10^{-10}}{10^8} \sim - 20s$ Which is the average time for an error to occur. The higher the PSNR, the lower would be the corresponding BER. Thus the noise is bit 1 and bit 0 is the same and in such the optimum setting of the threshold value is at the midpoint of the one and zero levels and the BER is related to PSNR through the following equation:

$$BER = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\sqrt{PSNR}}{\sqrt{2}} \right) \right] \quad (4)$$

Where erf represent the error function

3.7.3 Peak Signal to Noise Ratio: The peak signal to noise ratio (PSNR) are used to compare quality of compression image. The MSE represents the aggregate squared error between the compressed and the original

image while, PSNR represents a measure of the peak error. The lower the value of MSE, the lower the error of the embedded image. To calculate the PSNR, we first compute the MSE using the following equation:

$$MSE = \frac{\sum_{P,Q} [O_i(p,q) - W_i(p,q)]}{P * Q} \quad (5)$$

Where O_i is the cover image and W_i water-marked image. P and Q are the number of rows and columns in the input images. The block computes the PSNR using the following:

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right) \quad (6)$$

Where R is the maximum fluctuation in the in-put image datatype. For example, if the image has a double precision floating point datatype, then R is 1. If it has an 8-bit unsigned integer datatype, R is 255 and so forth. In general when the PSNR is 35dB or larger, then the two images are virtually indistinguishable by human observers. Image quality depends upon attack, watermark length and size of image.

3.7.4 Structural Similarity Index: This method gives similarity measure between two images. SSIM can take a value in the range -1 to 1, where a value of 1 means that the two images are completely similar to each other. The SSIM is computed using the following equation in (7):

$$SSIM = (I_i, I_w) = \frac{(2\mu_i\mu_{I_w} + C_1) * (2 * Cov + C_2)}{(\mu^2 I + \mu^2 I_w + C_1) + (\sigma^2 i + \sigma^2 I + C_2)} \quad (7)$$

Where $\{C_1 = (k_1 L)^2; k_1 = 0.01; C_2 = (k_2 L)^2; k_2 = 0.03\}$ and μ_i and μ_{I_w} the average of I and I_w respectively. Cov is the covariance of I_w , C_1 and C_2 are variable to stabilize the division with weak denominator, and L is the dynamic range of pixel values given by

$$L = 2^{number\ of\ bits\ per\ pixel} - 1$$

3.8 Applicability and Parameters used in GP

In this section, the various parameters associated with GP are discussed. The GP techniques applicable in the watermarking technique of IP that are presented in this section in terms of their domains but however, share some similarities in solving problems related to other domains.

3.8.1 Representation: In most of the approaches, GP individuals are represented as tree structure. In this thesis a hierarchical representation for the GP is illustrated (see figure 14). This was also represented by [4], in which they use an intelligent and reversible watermarking technique to construct the GP expression tree.

3.8.2 Function Set: Function set was chosen according to the problem in hand. The function set consists of the functions of the programs. The functions are several mathematical functions such as addition, subtraction, division, multiplication and other more complex functions. For instance for regression related problem the function set comprises of arithmetic operations (*, +, /, -, etc.).

3.8.3 Terminal Set: The terminal set similarly does not have any specific predefined set. The GP terminal set is comprised of variables also called program input, constants or random input. Some of these variables could be the velocities and accelerations of the gun, bullet and target.

3.8.4 Fitness Function: The most challenging and most important concept of genetic programming is the fitness function. The fitness function determines how well a program is able to solve the problem. For example in this research a fitness function was used to evaluate the performance of each candidate expression and determine whether to keep the expression for the next generation or ignore it. Additionally fitness function measures how good and bad is a specific region within the search space. Different fitness functions depending on the nature of the issue, have been utilized as an evaluation measure during the search process namely root-mean-square-error (RMSR), PSNR, accuracy and so on.

3.8.5 Initial Population: If a priori knowledge about the features of the desired solutions is not known, then the initial individuals are generated randomly. Generates a population of points at each iteration. The best point in the population approaches an optimal solution. Select the next population by computation which uses random number generators given by the equation:

$$g = \sin(\alpha x) * \sin(\alpha y) e^{\frac{-(x+y)}{\sigma}} \quad (8)$$

3.8.6 Selection Method: In GP evolution cycles mainly there are two types of selection methods used namely: parent selection and survivor selection. For the purpose of this research tournament selection was used since it's the most widely used selection mechanism. In the parent selection individuals performed an individuals which are produced from selected parents.

3.8.7 Genetic Operators: In order to intro-duce diversity among the individuals of the population, different genetic operators namely: crossover, mutation and reproduction are used for the generation of an o spring and thus explained below:

3.8.7.1 Crossover: Two primary operations exist for modifying structures in genetic programming. The most important one is the crossover operation. In the crossover operation, two solutions are sexually combined to form two new solutions or off- spring. The parents are chosen from the population by a function of the fitness of the solutions.

3.8.7.2 Mutation: Mutation is another important feature of genetic programming. Two types of mutations are possible. In the first kind a function can only replace a function or a terminal can only replace a terminal. In the second kind an entire subtree can re- place another subtree. Figure 14 explains the general structure and concept of genetic programming.

3.8.7.3 Reproduction: The reproduction function determines how to expand the population based on the existing candidates. Modifying the behavior and hyperparameters of the reproduction function is one of the most complex and impactful parts of creating a genetic programming, as the reproduction function is what determines how the population changes over time.

4. EXPERIMENTAL RESULTS

In this section, we present quality measure matrix of our watermarking scheme. The algorithm have been evaluated by various simulations exercises carried out in both medical images and other standard images such as Lena, Baboon, Pepper, Barbara and so on and are used during the experimentation of our schemes. Some of the images we tested are of standard size 512x512 and results are shown in figure 7. The two main evaluation parameters use to check for image imperceptivity are PSNR and SSIM. The results are recorded in figure 7. Imperceptivity and robustness describes the measure of visual perception to the images after watermarked has been conducted when compared to original image. On the other hand, robustness indicates the resistance of the algorithm when it undergone various attacks. These exercise have been demonstrated and recorded against state- of-the-art watermarking techniques. NCC and BER have been tested and compared with other techniques shows it resistance to various attacks as shown in Figure 8 and 9 of our scheme. Therefore, from figure 10 it could be noted that the proposed watermarking scheme I provides better imperceptivity than the schemes under review and the watermark is retrieved without errors at the receiving side. Similarly, figure 11 shows PSNR comparative analysis for both scheme I and scheme II for gray scale images. Scheme II shows better performance than scheme I due to it intelligent GP generation technique. It has also been indicated that the closer NCC value to unity the more robust the algorithm and capability to resist signal processing attacks.

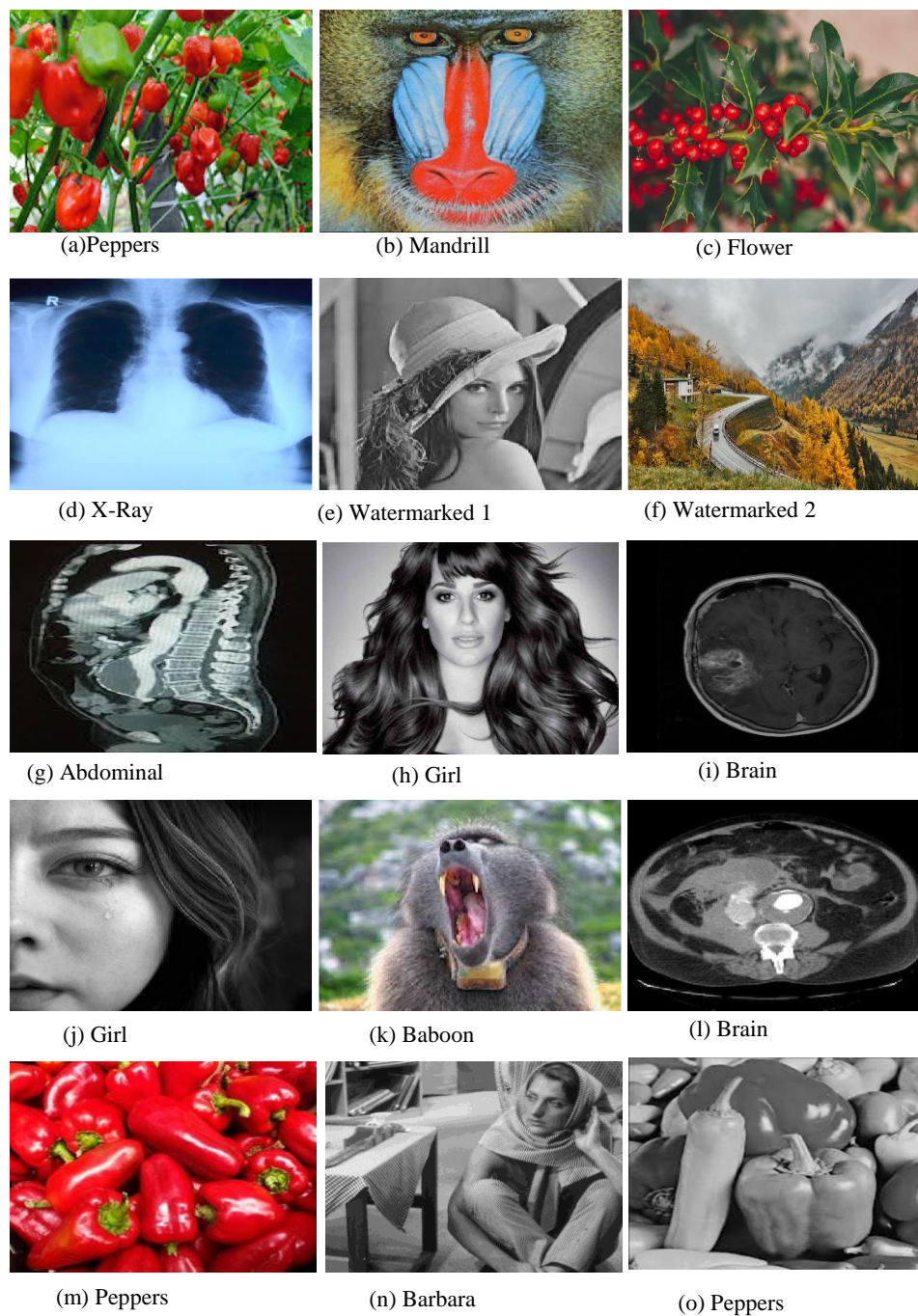


Figure 7: Some of the gray scale, color images and the watermarks used in the proposed water-marking scheme(s)

4.1 Imperceptivity Analysis

In this section, we present perceptual quality of the watermarked images from our simulation exercise and thorough analyze are presented for evaluation. We use two parameters to evaluate the measure used for image imperceptivity such as PSNR and SSIM.

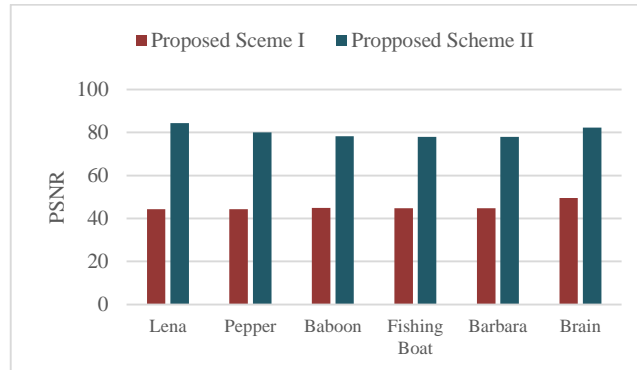


Figure 11: Scheme I Vs Scheme II Imperceptivity comparison with different gray scale images

4.2 Robustness Analysis for Fragility

In order to determine the robustness of the watermarking technique we presents two key parameters NCC and BER analysis. For copyright protection perceptiveness and robustness are key factors for any watermarking algorithm. In this section, many watermarked images are been subjected to attacks and the result are recorded and presented. Attacks such as filtering, noise addition and compression has been undertaken on the watermarked image. Individually for the images under observations were tested and as well as simultaneous attacks. The objective of NCC and BER were compared against existing state-of-the-art outcomes and were recorded as figure 6, 7 and table 2.1 respectively. The results shows high effectiveness except for NCC which is more robust then our scheme.

Table 2: BER (%) average comparison for various attacks

Attack Type	Hurrah et al. 2019	Proposed Scheme I
None	0	0
JPEG/DF=60	0	0
JPEG/DF=30	0	0
JPEG/DF=20	0.30	0.50
JPEG2000/CR=4	0	0
JPEG2000/CR=8	0.39	0.25
Gaussian Noise @0.01	0.68	0.52
Salt and Pepper Noise@1%	8.05	7.45
3x3 Median Filter	5.98	5.30
3x3 Gaussian LPF	0.04	0.02

4.3 Scheme I Robustness Analysis

Scheme I presents embedding a robust water- mark in both gray scale images and color images. The quality measurement of the schemes are presented in figure 8 and 9 which were subjected into various attacks. The analysis of scheme I proves that watermarked algorithm is robust against different kind of attacks carried out in mage watermarked Lena. The algorithm shows robustness to Gaussian low pass filtering whilst BER values of 0.2 to 5 percent could be seen for different attacks. For low quality factor of 25, our schemes only report 0.1 percent of BER.

The proposed scheme I has been compared for JPEG compression with [10] and the result has been presented in table 2.




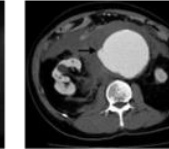
Watermarked Image				
PSNR (dB)	44.67	44.69	44.7	44.74
SSIM	1	1	1	1
Extracted Watermarked	Copyright	Copyright	Copyright	Copyright
BER (%)	0	0	0	0
NCC	0.0038	0.0038	0.0038	0.0038

Figure 8: Watermarked for gray scale images, extracted watermarks and corresponding parameter factors of = 36




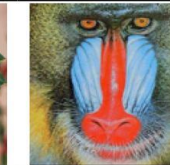
Watermarked Image				
PSNR (dB)	44.63	44.63	44.30	44.96
SSIM	1	1	1	1
Extracted Watermarked	Copyright	Copyright	Copyright	Copyright
BER (%)	0	0	0	0
NCC	0.0039	0.0039	0.0039	0.0039

Figure 9: Watermarked for color images, extracted watermarks and corresponding parameter factors of = 36

4.4 Comparison of Scheme II with existing Techniques

Coming up with better or suitable techniques has to do with maximum literature review and finding the gap of already state-of-the-art watermarking techniques ([29],[28],[24],[3],[21],[4]), with the proposed scheme II. Arsalan's technique is interesting but its disadvantage of fragility, the extraction processing will be not reversible when any filter is applied. As a result the image is degraded when there is no suitable space for embedding the map location. Henceforth, a suitable solution is needed to address image perceptivity and capacity in order to generate move gap for the overhead. Arsalan [4] presented that evolve a mathematical function that selects the coefficient for companding with a maximum payload of 0.7bpp but however, the required companding less or greater than the threshold producing degradation in an image but better than the approach used by [29]. Hence the existing reversible watermarking schemes provide some tradeoff between robustness and imperceptivity of watermarked. Therefore, we present a GP based algorithm of scheme II to improve the margin that depends of image imperceptivity and robustness. The proposed scheme II recorded a minimum of PSNR 58.89dB and SSIM 1 when measured in terms of imperceptivity (see figure 5). Again for NCC value of 1 and BER of 0 when measured for robustness. Our scheme II outperform better than the approach presented by [4]. In our scheme II, various grayscale images were subjected to simulation and figure 6 for Lena image the compromise between payload and imperceptivity was displayed. Our proposed scheme II shows better improvement of PSNR and SSIM when compared to [4] results with maximum payload (0.1 to 0.7). It thus indicates (figure 6) that at low payload value the PSNR and SSIM show high perceptivity of the watermarked images. This is because the GP has to produce a generic expression applicable to all the blocks in the image.


Watermarked Image				
PSNR (dB)	58.89	86.69	96.61	82.05
SSIM	1	1	1	1
Extracted Watermarked				
BER (%)	0	0	0	0
NCC	1	1	1	1

Figure 5: Watermarked images, extracted watermarked and corresponding parameter at gain factor = 50

4.5 Scheme II Imperceptibility Analysis

As stated in scheme I algorithm the parameters used to measure imperceptibility of watermarked base algorithm are the values of PSNR and SSIM. Using standard images of 512x512 and 256x256 for various images scheme II provides higher performance when compared to [4] approach as shown in table 3 and table 4. Scheme II utilizes three out of four sub-bands for water- marking embedding and frequency sub-band of 8x8 DCT in order to maximize payload. The outcomes of scheme II testing for both medical images and other standard images have been accomplished by reversing the payload of 0.6bpp. The complexity of the GP algorithm is shown in figure. These results are obtained on MAT- LAB 2019a on Intel R Core™ i7 -33770 CPU @ 2.30GHZ with RAM of 16GB. Our approach outperform the [4] approach.

Table 3: PSNR (dB) Vs payload (bpp) based comparison of proposed scheme II with against [4] approach

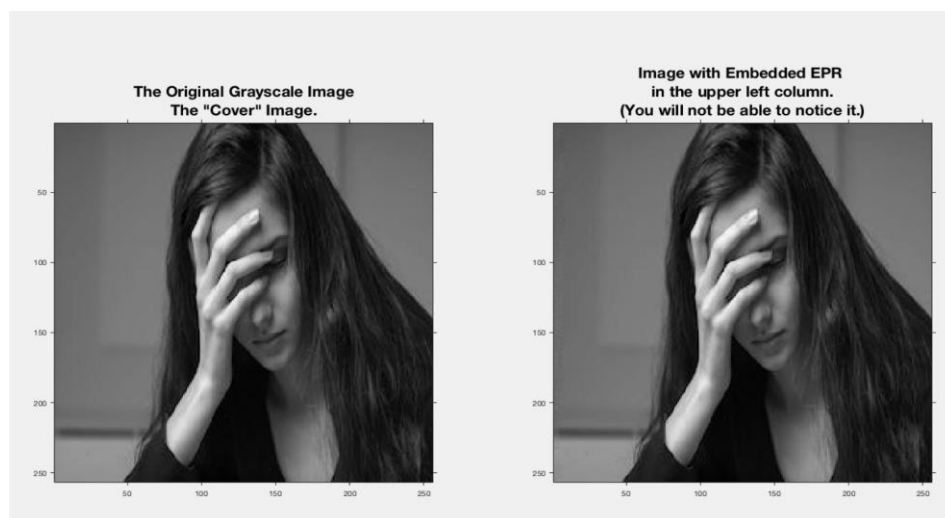
Image	Payload (bpp)	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Lena	Effective payload	0.58						
	Arsalan et al. [4]	50.58	46.71	45.20	44.30	43.60	42.60	41.87
	Proposed scheme II	58.89	53.74	52.23	51.33	50.63	49.3	48.30
Barbara	Effective payload	0.48						
	Arsalan et al. [4]	50.24	47.3	45.68	44.81	44.11	43.10	42.43
	Proposed scheme II	91.15	88.21	85.71	84.82	84.12	83.21	83.11
X-ray	Effective payload	>0.6						
	Arsalan et al. [4]	54.43	50.70	49.03	48.11	47.28	46.29	45.53
	Proposed scheme II	84.20	80.25	78.35	77.8	75.3	75.01	74.21
Belly	Effective payload	>0.6						
	Arsalan et al. [4]	52.12	48.94	47.37	46.55	45.83	44.93	44.20
	Proposed scheme II	78.30	76.30	75.23	72.30	70.3	68.23	67.85
Girl	Effective payload	>0.6						
	Arsalan et al. [4]	56.79	52.28	50.38	49.20	48.10	47.27	46.67
	Proposed scheme II	94.30	92.34	91.03	89.04	86.30	81.30	81.25

Table 4: Comparative display of SSIM proposed Scheme II technique against [4] payload

Image	Payload (bpp)	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Lena	Effective payload	0.58						
	Arsalan et al. [4]	0.9939	0.9886	0.9836	0.9794	0.9783	0.9707	0.9707
	Proposed scheme II	1	0.9947	0.9897	0.9855	0.9844	0.9768	0.9778
Barbara	Effective payload	0.48						
	Arsalan et al. 2017	0.9968	0.9943	0.9910	0.9890	0.9869	0.9832	0.9811
	Proposed scheme II	1	0.9975	0.9942	0.9922	0.9921	0.9878	0.9858
X-ray	Effective payload	>0.6						
	Arsalan et al. [4]	0.9971	0.9932	0.9903	0.9878	0.9853	0.9817	0.9783
	Proposed scheme II	1	0.9955	0.9940	0.9922	0.9920	0.9912	0.9873
Belly	Effective payload	>0.6						
	Arsalan et al. [4]	0.9961	0.9927	0.9891	0.9872	0.9848	0.9853	0.9836
	Proposed scheme II	1	0.9968	0.9959	0.9938	0.9923	0.9911	0.9875
Girl	Effective payload	>0.6						
	Arsalan et al. [4]	0.9983	0.9958	0.9936	0.9917	0.9896	0.9872	0.9854
	Proposed scheme II	1	0.9982	0.9973	0.9968	0.9957	0.9940	0.9922

4.6 Embedding and extraction of EPR

This section, we present an embedded image along with data (in string) to be embedded or hidden in the image. The EPR may include strings of data such as PID, Name, Age, Blood Pressure (BP), Sugar level, Proposed Treatment and other related medical information for the patient are embedded in the image. Scheme II also produces indistinguishable images of both when apparently placed side by side. The process was achieved by providing 8-bit plane where a user is required to insert the bit frequency bit plane. At 1-5 bits provides better performance than 6-8bits (figure 13).

**Figure 13:** EPR Embedded and Restored Method

4.7 Analysis of GP Watermark Scheme II

Figure 14 (c) shows the GP tree generated for a suitable expression of 50 generations. The graph prove the various of the structural complexity of the algorithm. Again, genetic operators computation, population diversity, fitness function, accuracy versus complexity and Pareto front charts are generated during the computation process. It can be learnt that from these figures 14 (a-f) after few generations learning ability of GP improves and produces a suitable expression for any given PSNR value and improve generation by generation which serve as the basis of Darwin evolutionary theory.

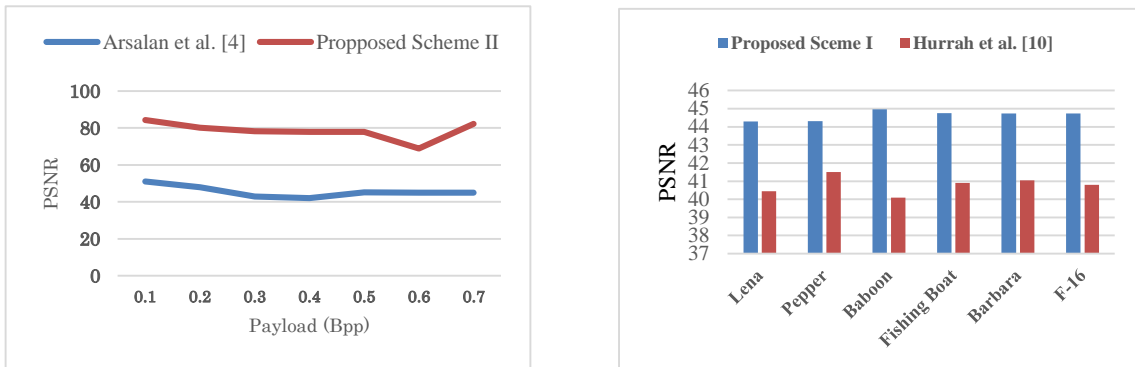


Figure 6: Comparison between proposed scheme II for effective payload (0.1 to 0.7)

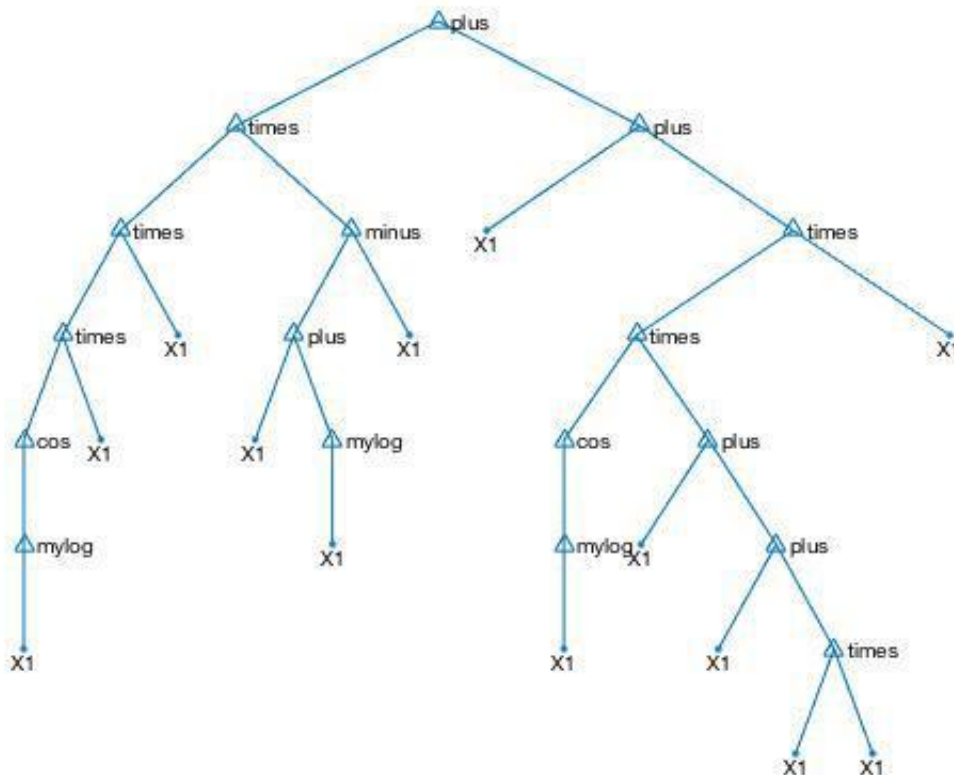


Figure 14: GP tree generated during computational analysis

4.7.1 GP Population Diversity Vs Fitness Function

From the population diversity sinusoidal graph line shows 84 unique generation (line with green). For population of 84 against 50 generations of the GP during the testing phase of the algorithm. The result shows (see population diversity figure 5.10a) within 10-15 generations records highest population diversity of 93.5 and between 35-50 generations the population diversity shows an optimum of 90. In the other hand for fitness function, we records seven different parameter lines. The horizontal blue line shows a maximum of 0.87485 log10 fitness against 50 generations, green line shows the median with 1.7383, yellow line indicates average and the two broken gray lines shows a negative average of -0.13542 standard deviation and a positive 12.044 average standard deviation respectively. The line with dark blue shows the best log10 fitness function of 0.87075 against 50 generations and red line shows the test fitness of 16.8959 against 50 generations (see figure 15b). The algorithm performances relatively better over 50 generations with high capacity of embedding EPRs for the healthcare domain. At this stage, since the GP has computed the population diversity against the specified 50 generations it will now enhance the capacity of the algorithm to decide the expression of the GP. Again, having computed the population diversity next is to determine the fitness values for all the possible generations that survives during the crossover evolutionary process. The GP scheme II shows that for optimum value of the population diversity better expression values could be decided for future generations and will provide better generational values during the process.

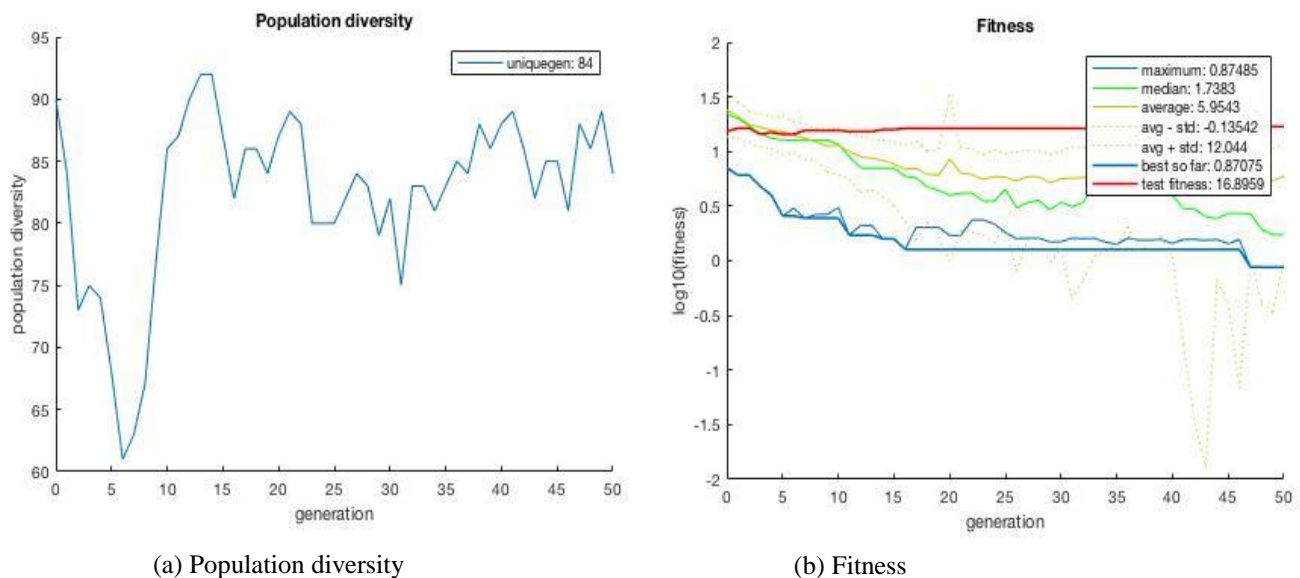


Figure 15: GP Population diversity Vs Fitness function

4.7.2 GP Accuracy versus Complexity and Pareto front

There are three key concepts used to develop GP accuracy namely; yellow line shows the number of nodes during the GP generations of maximum of 40 against 50 generations. The result shows that between 0-10 the number of nodes reaches a maximum of 40 and drops steadily and increases again forming a sinusoidal function and at 20 generations the number of nodes increases again between 35-38 and drops steadily at 30 against 50 generations (see figure 16b). The line with red indicates level of accuracy against 50 generations. The result reveals that at level of 10-12 records the highest and drops at 10 with steady values for 50 generations. Essentially, the blue line indicates the fitness and starts between 5 and 10 against 50 generations and drops slightly between 0 and 5. The charts show the accuracy versus complexity of the GP algorithm generations overtime and over subsequent generations (see figure 16a). A Pareto front chart was also generated during the testing phase, the blue line shows the best number of nodes utilized (fitness against nodes). The green stars show the current population of 70 with randomly scattered. The pink line shows the test fitness against the nodes and red line determine the Pareto front against the best fitness among nodes, current population and the test fitness. The

pareto front function is shown on figure 16b) of the pareto front chart. It could be noted that for successive generations the the accuracy versus complexity increases after each generations and improve the accuracy for the GP and the structural complexity becomes more complex. Hence making the algorithm more interesting and development and implementation into cloud database servers using the proposed architecture.

In the other hand for the pareto front deterministic value could be done between fitness against nodes for every successive generations and increases steadily as we go across the interface between nodes for best fit, current population generations and test fitness. Looking at the diversity of the population distributions (see figure 16b), it was found that at higher fitness values the current population generation are dispersed and with high concentration of nodes against the node axis. The re-researcher in his wisdom, thought of an adequate technique to improve the GP as a fundamental concept to enhance security in the Cloud by coming with scheme II proposed technique of watermarking intelligent. The technique is not only computing the parameters of the watermarked but also provide an intelligent method to enhance the efficiency of watermarking algorithm.

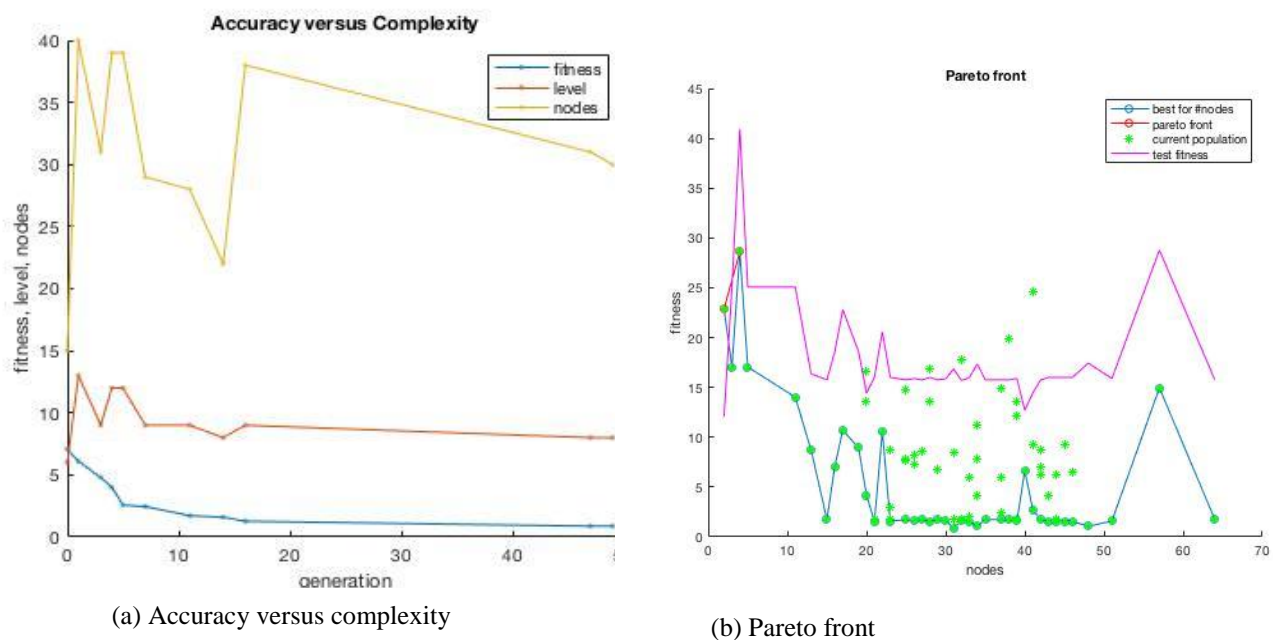


Figure 16: GP Accuracy versus complexity and Pareto front

4.7.3 GP Structural Complexity and Genetic Operators

Figure 17a of the GP generations shows the structural complexity of our algorithm. The tree depth of 10 , tree size and the percentage introns were plotted against 50 generations. The result proved that line with orange shows maximum size of 43, broken orange line with a star shows the best so far size of 30. Thin orange line indicates the average size of 34, pink broken line with star provides best so far introns and close to the bottom of the generation axis and the pink line shows average introns of 1427. Alternatively, blue broken lines with star reveals the depth best so far of 8 and line blue shows the average depth of 10.49 against 50 generations while black dot line provides the average tree fill of 34.1581. The result could be noted that the structural complexity of the GP generations of 50 provides a novelty of Scheme II approach with relatively high performance against other existing approaches.

Similarly, figure 17b shows the operator probability frequency against 50 generations of the GP algorithm. From the chart it can be found that the blue line indicates the probability that was obtained for crossover of 0.7421 indicating ex-pression of the GP for this scheme. In the other hand the blue line with x across the line shows the number of done crossover during the computational process (12). Orange line proves the probability of mutation reaching a 0.2579 over 50 generations and the orange line with x crossing the line proves the done mutations (0). Con-sequently, the dotted lines shows the cumulative frequency crossover of 2209 and

cumulative mutation of 317 over 50 generations hence the results shows that our scheme provides a novelty of GP generation over successive generations. And the black line indicates the reproductions during the genetic programming generations (5). Our results shows that at higher probability of 0.9 or 1 provides certainty for crossover to take effect (blue line) and the number of mutations drops from 0.5 to 0 probability of possible mutations to occur.

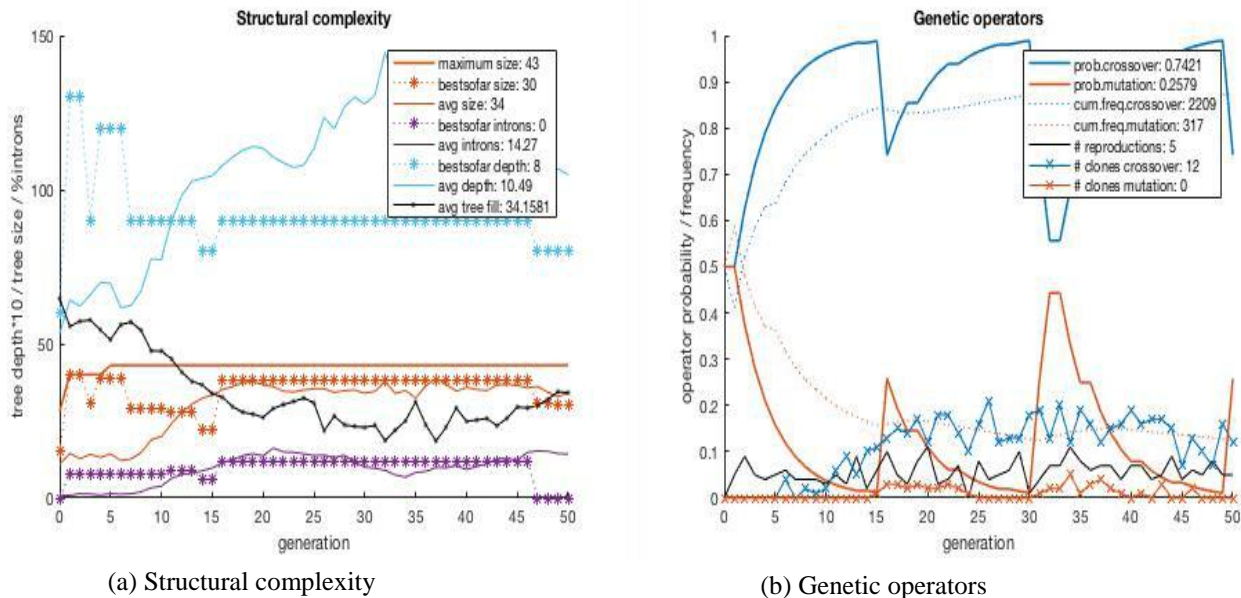


Figure 17: GP structural complexity and computation of genetic operators

4.7.4 Desired versus obtained

Again figure 18 shows the desired versus obtained generation for 50 successive generations. For subsequent generations, the GP keeps the best generation to be used for crossover and reproduction. The results reveals that the best generation results that was obtained during the testing phase of 50 generations are 0, 1, 3, 4, 5, 7, 11, 14, 16, 47, and 49 respectively. This results provide not only a novelty for the algorithm but also provides a theory of the evolutionary theorem postulated by Charles Darwin.

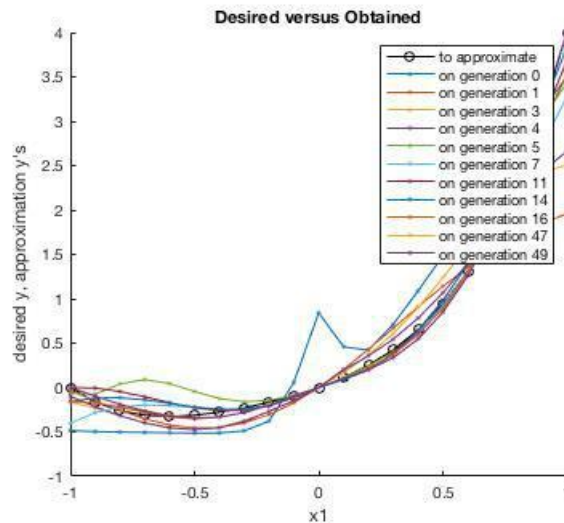


Figure 18: Desired versus obtained

5. CONCLUSION

This paper present two watermarking schemes namely scheme I and scheme II respectively and has been proposed for copyright protection, content protection and authentication of digital media through content exchange in an insecure broadcast in the cloud. The schemes have been evaluated for both grayscale images and color images and some standard medical images. Scheme I used combined features of DWT and DCT for embedding of the watermark for both color and grayscale images. Scheme II embeds a GP and DCT or DWT base algorithm for protection of patient related information in the healthcare domain. EPR information was also used through steganography technique to hide the EPR in the image. Both Scheme I and Scheme II provide robustness and imperceptivity performance when compared to existing watermark technique of [10] and [4] respectively. It has also been observed that at a given effective payload our scheme II improves imperceptivity for an image comparison to [4] and our GP based technique able to compute the structural complexity (see figure). For scheme I, we intended to extend the work for facilitate more improvement of robustness and imperceptivity of the image and hoped to be tested in video watermarking techniques.

6. REFERENCES

- [1] Micheal Agbaje, Oludele Awodele, and Chibueze Ogbonna. "Applications of digital water-marking to cyber security (cyber watermarking)". In: Proceedings of Informing Science & IT Education Conference (InSITE). 2015, pp. 1-11.
- [2] Lingling An et al. "Robust lossless data hiding using clustering and statistical quantity histogram". In: Neurocomputing 77.1 (2012), pp. 1-11.
- [3] Muhammad Arsalan, Sana Ambreen Malik, and Asifullah Khan. "Intelligent reversible water-marking in integer wavelet domain for medical images". In: Journal of Systems and Software 85.4 (2012), pp. 883-894.
- [4] Muhammad Arsalan et al. "Protection of medical images and patient related information in healthcare: Using an intelligent and reversible watermarking technique". In: Applied Soft Computing 51 (2017), pp. 168-179.
- [5] Walter Bender et al. "Techniques for data hiding". In: IBM systems journal 35.3.4 (1996), pp. 313{336.
- [6] Gaurav Bhatnagar and QM Jonathan Wu. "Biometrics inspired watermarking based on a fractional dual tree complex wavelet transform". In: Future Generation Computer Systems 29.1 (2013), pp. 182-195.
- [7] Abbas Cheddad et al. "Digital image steganography: Survey and analysis of current methods". In: Signal processing 90.3 (2010), pp. 727-752.
- [8] Ahmed M Elmisery, Seungmin Rho, and Dmitri Botvich. "A distributed collaborative plat-form for personal health pro les in patient-driven health social network". In: International Journal of Distributed Sensor Networks 11.9 (2015), p. 406940.
- [9] Aggeliki Giakoumaki, Sotiris Pavlopoulos, and Dimitris Koutsouris. "Secure and efficient health data management through multiple watermarking on medical images". In: Medical and Biological Engineering and Computing 44.8 (2006), p. 619.
- [10] Nasir N Hurrah et al. "Dual watermarking framework for privacy protection and content authentication of multimedia". In: Future Generation Computer Systems 94 (2019), pp. 654{ 673.
- [11] HIPAA Journal. 2018 Was a Record-Breaking Year for Healthcare Data Breaches. <https://www.hipaajournal.com/analysis-of-healthcare-data-breaches>. 2019.
- [12] A Kannammal and S Subha Rani. "Two level security for medical images using watermarking/encryption algorithms". In: International Journal of Imaging Systems and Technology 24.1 (2014), pp. 111-120.

- [13] Basant Kumar et al. "Secure spread-spectrum watermarking for telemedicine applications". In: Journal of Information Security 2.02 (2011), p. 91.
- [14] Ram Kumar et al. "Histogram thresholding in image segmentation: A joint level set method and lattice boltzmann method based approach". In: Information Technology and Intelligent Transportation Systems. Springer, 2017, pp. 529-539.
- [15] Hussain Nyeem, Wageeh Boles, and Colin Boyd. "A review of medical image watermarking requirements for teleradiology". In: Journal of digital imaging 26.2 (2013), pp. 326-343.
- [16] Geovandro CCF Pereira et al. "Performance evaluation of cryptographic algorithms over IoT platforms and operating systems". In: Security and Communication Networks 2017 (2017).
- [17] R Lakshmi Priya and V Sadasivam. "A SURVEY ON WATERMARKING TECHNIQUES, REQUIREMENTS, APPLICATIONS FOR MEDICAL IMAGES." In: Journal of Theoretical & Applied Information Technology 65.1 (2014).
- [18] Niels Provos and Peter Honeyman. "Hide and seek: An introduction to steganography". In: IEEE security & privacy 1.3 (2003), pp. 32-44.
- [19] Raul Rodriguez-Colin, Feregrino-Urbe Claudia, and Gershom de J Trinidad-Blas. "Data hiding scheme for medical images". In: 17th International Conference on Electronics, Communications and Computers (CONIELECOMP'07). IEEE. 2007, pp. 32-32.
- [20] Abhilasha Sharma, Amit Kumar Singh, and Satya Prakash Ghrera. "Robust and secure multiple watermarking for medical images". In: Wireless Personal Communications 92.4 (2017), pp. 1611-1624.
- [21] A Siddiqi and A Khan. "High capacity reversible image watermarking using error expansion and context-dependent embedding". In: Electronics Letters 51.13 (2015), pp. 985-987.
- [22] Amit Kumar Singh et al. "Multiple watermarking on medical images using selective discrete wavelet transform coefficients". In: Journal of Medical Imaging and Health Informatics 5.3 (2015), pp. 607-614.
- [23] Prabhishek Singh and RS Chadha. "A survey of digital watermarking techniques, applications and attacks". In: International Journal of Engineering and Innovative Technology (IJEIT) 2.9 (2013), pp. 165-175.
- [24] Jun Tian. "Reversible watermarking by difference expansion". In: Proceedings of workshop on multimedia and security. Vol. 19. 2002.
- [25] Blesson Varghese and Rajkumar Buyya. "Next generation cloud computing: New trends and research directions". In: Future Generation Computer Systems 79 (2018), pp. 849-861.
- [26] Li Weng, Laurent Amsaleg, and Teddy Furon. "Privacy-preserving outsourced media search". In: IEEE Transactions on Knowledge and Data Engineering 28.10 (2016), pp. 2738-2751.
- [27] Zhihua Xia et al. "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing". In: IEEE Transactions on Information Forensics and Security 11.11 (2016), pp. 2594-2608.

- [28] Guorong Xuan et al. "Distortionless data hiding based on integer wavelet transform". In: Electronics Letters 38.25 (2002), pp. 1646-1648.
- [29] Guorong Xuan et al. "Reversible data hiding using integer wavelet transform and companding technique". In: International Workshop on Digital Watermarking. Springer, 2004, pp. 115-124.
- [30] Zheng Yan et al. "Encrypted data management with deduplication in cloud computing". In: IEEE Cloud Computing 3.2 (2016), pp. 28-35.
- [31] Aditi Zear, Amit Kumar Singh, and Pardeep Kumar. "A proposed secure multiple watermarking technique based on DWT, DCT and SVD for application in medicine". In: Multimedia tools and applications 77.4 (2018), pp. 4863-4882.
- [32] Sheng Zhong et al. Security and Privacy for Next-Generation Wireless Networks. Springer, 2019.