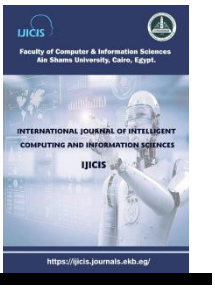
IJICIS, Vol.22, No.1, 76-87
DOI: 10.21608/ijicis.2021.82820.1108

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/


IJICIS
Faculty of Computer & Information Sciences
Ain Shams University, Cairo, Egypt.
INTERNATIONAL JOURNAL OF INTELLIGENT COMPUTING AND INFORMATION SCIENCES
IJICIS
https://ijicis.journals.ekb.eg/

Protein Key Generation for Secure CT-Chest Images Encryption

Sara A.Shehab*

Computer Science Dep, Faculty of Computer and Artificial Intelligence, Sadat city, Egypt

Sara.shehab@fcai.usc.edu.eg

Received 2021- 06-27; Revised 2021-10-18; Accepted 2021-11-02

Abstract: *Nowadays, one of the most complex problems in telemedicine and E-health is the preservation of patient data due to the integration between the development of technology and the medical sector. To protect patient privacy, the transmission of the secured medical image requires adequate techniques. This study aims at encrypting COVID-19 images of Computed Tomography (CT) chest scan into secured and sensitive cipher images for the infected patient. To achieve a high degree of security in the encryption process, protein key generation for the encryption process has been proposed. This study aims to encrypt images using 2 round AES plus Protein key. The histogram has been used to estimate the degree of security for the proposed method. Four criteria have been selected to evaluate the degree of security for the proposed method Number of Pixel Change Rate, Correlation coefficient, Entropy, and Unified Average Changing Intensity. The result indicated that the proposed method has 99.5% and above NPCR, Correlation coefficient close to zero, UACI above 30%, and Entropy near to 8. The results confirm that the proposed method achieves a high level of security and sensitivity when compared with previous work. Therefore, the proposed method can be considered as a successfully applied algorithm to satisfy the security requirements of transmitting CT images for COVID-19 patients.*

Keywords: *Protein, Number of Pixel Change Rate, Histogram, Entropy, Unified Average Changing Intensity*

1. Introduction

Cryptography is a process of encoding text or images with a high level of security to make them unreadable and protected [1]. Cryptanalysis is defined as a technique of converting unreadable text into readable one without loss of information. The combination between cryptography and cryptanalysis is

* Corresponding author: Sara A.Shehab
Computer Science Dep, Faculty of Computer and Artificial Intelligence, Sadat city, Egypt
E-mail address: Sara.shehab@fcai.usc.edu.eg

defined as Cryptology. Many cryptology techniques are complex before the age of computers. Nowadays, the generation of computers makes complex cryptographic algorithms easy to use. The encoding process means converting the plain text into ciphertext. Whereas Decryption is the process of encoding cipher text to plain text [2]. The used algorithm and key in this algorithm are the two main key parameters when applying the Encryption and Decryption process. When the key used in the encryption process same as the key used in the decryption process, the key is defined as a symmetric key. Whereas when the key differs in two processes it is defined as an asymmetric key. Key is considered as the most important attribute in security achievement. Some of the most used security algorithms are the international data encryption algorithm (IDEA) [3, 4], data encryption standard (DES), blowfish, and advanced encryption standard (AES) [5, 6].

AES is one of the asymmetric algorithm keys where the encryption process uses the same key decryption process used. One of the most important key parameters in AES is the substitution network. Unlike DES, the Feistel network can’t be used in AES. When processed AES algorithm 128-bit block with the fixed size is used, and key size of 128,192 and 256 bits is used. AES works with a 4 × 4 array matrix of bytes. AES specifies the number of rounds to convert a plain text to ciphertext. If the block size equals 128 bits, then it uses ten rounds, whereas if the block size equals 192 bit it uses 12 rounds, and if the block size equals 256 bit it uses 14 rounds. AES steps are discussed in Figure 1 [7].

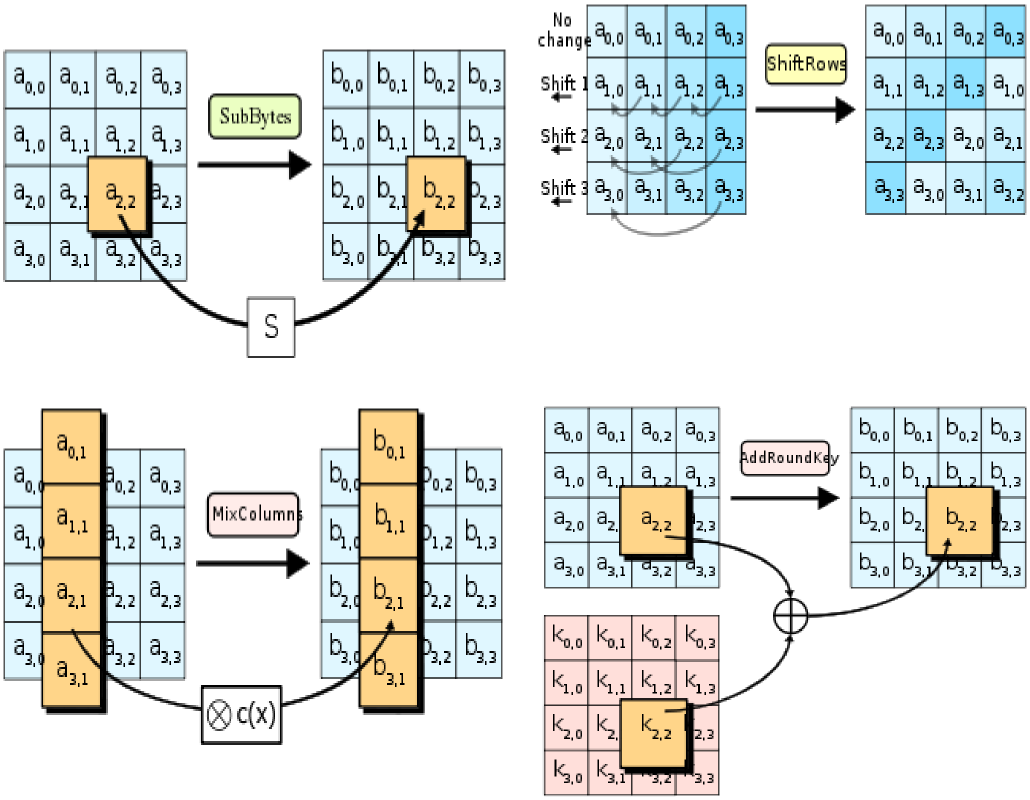


Figure 1:AES Steps

DNA (deoxyribonucleic acid) has a structure of two complementary strands bonds in a double helix to form the final structure. DNA consists of four bases Cytosine (C), Adenine (A), Thymine (T), and Guanine (G). Due to the high level of security for data hiding and easily generate symmetric key DNA cryptography was proposed. There are many advantages to select encryption with DNA. Firstly, DNA structure helps the computer to compute 100 million instructions per second. Secondly, to store one bit, DNA requires 1bit per cubic nanometer whereas the media storage requires from 10 to 12 cubic nanometer. Finally, DNA doesn't need a power requirement as it is based on its structure and chemical bonds. In DNA bases, 1021 bases equal 1 gram of DNA and 1 gram equal to data with 108 terabytes size, so DNA can use few grams to store all data in the world. DNA can easily be converted to RNA by changing the base Thymine (T) to (U) Uracil. After that, the protein codon will be performed from each three RNA bases.

2. DNA cryptography

Biological structures can help in securing information that is defined as DNA computing. In 1994, Leonard was the first person to propose working with DNA computing [8]. He uses it in solving complex algorithmic problems, now data can be transmitted and stored using DNA. Cryptography means using DNA sequences to hide data in it. DNA Sequences are the most commonly used in all work in DNA cryptography when trying to encrypt binary codes to other forms. In the future, securing data with DNA computing will become one of the most important issues. There are some advantages of DNA computing: -

1. Minimum storage requirement, when storing 108 terabytes of data it will need only 1 gram of DNA in calculation whereas a compact volume can store a large amount of data.
2. Speed improved; data stored in DNA computing techniques can operate 100 times faster than nowadays computers.
3. Low power requirement, when compared with modern computers, DNA computing needs no power at all.

3. Related Work

In 2016, T.mahalaxmi proposes a symmetric algorithm with DNA cryptography that achieves a secure data transfer mechanism. Firstly, the data (text or image) was inputted after that it was converted to its equivalent ASCII value, then convert to binary equivalent. Secondly, the binary values were encoded to DNA bases. A private key is randomly selected and converted to ASCII. Finally, the DNA code used a private key for encrypting the data and clinical permutation. The proposed algorithm was implemented with java and listed as a modern symmetric key encryption technique. This method uses a much faster-encoded technique [9].

In 2016, N.gulati proposes a new technique using XOR operation with a one-time pad (OTP) scheme and DNA cryptography. XOR operation is applied between OTP and binary data. A suitable method with some precondition that applied to the OTP scheme. The binary digits are converted into DNA sequences where T equals 01, A equals 00, C equals 10, and G equals 11. After that DNA bases are complemented, and the sequence is reversed from right to left. The result is encrypted data that send to the recipient. This algorithm achieves three levels of security, arithmetic XOR operation, OPT, and DNA complementary rule. The proposed method is simple since there are some preconditions applied to OPT that the user must take care of it [10].

In 2016, Sravanan used the DNA encryption and decryption technique and proposes an algorithm depends on to modify the secret algorithm Shamir proposed. Instead of a single user, a group is involved at the receiver end. There are some added security in-crop orated in the algorithm. The secret message can be encoded only when the decryption process includes all the clients in the group. To convert the message to ASCII values mathematical operations are performed. Then it is converted to DNA bases. DNA coding is used to decrypt the message after the message is transmitted to all clients in the group this step increases the security transmission for the multi-task application. This method is implemented with java and python. One disadvantages of this method is that method suitable only for a group of clients and the message cannot be decrypted if someone is missing from this group [11].

In 2016, E.S.Babu proposes a DNA cryptography method with biotic pseudo. To improve security, the splicing system methodology is used. It is randomly generated the key for this reason it is difficult to re-cipher the message. The robustness of the proposed method approved that it is more secured than common methods [12].

In 2017, Paspula proposes a new ciphertext generation procedure with a unique key generation method. It proposes two rounds to generate the key. First-round, an intermediate form of ciphertext has been generated by conventional cryptographic techniques whereas, in the second round, the intermediate form of ciphertext is encoded into final ciphertext. To confuse an intruder, fake DNA sequences were generated. One of the most disadvantages of this method that it increases the time and space complexity [13].

In 2020, Pramod proposed anew cryptosystem based on DNA with finite automata theory. The algorithm based on three parameters key generator, sender, and receiver. The sender generates a 256-bit DNA based secret key based on the attributes of the receiver, in the encryption process this key is used. Then, for coding the DNA sequence a randomly generated Mealy machine is used, which increases the level of security. The proposed scheme can protect the system against numerous security attacks, such as brute force attack, known plaintext attack [22].

In the related work, all related work encrypted data using DNA bases. the encrypted data achieve high level of security. But nowadays the DNA key generation can be expected. To increase the degree of security protein key proposed that can’t be expected.

4. Proposed Algorithm

Generating a key for encryption and the same for the decryption process has been proposed. Firstly, the user inputs the string value from the console, after that, it will be converted to its equivalent binary 7-bit, then encoding each two-bit to DNA-bases i.e. A=00, G=01, C=10, and T=11 (Table1). The next step is to get complimentary of the DNA bases i.e., complimentary of A=T, C=G (Table 2) and added to original bases. In the process of converting DNA to RNA only one base Thymine (T) will be changed to Uracil (U). Every three bases from RNA encode one protein. The method of generating protein key stopped at codon UAA, UGA, and UAG, the length between this codon is estimated and returns the longest length as a protein key. Finally, the protein key converted to binary value and used in two processes encryption and decryption. Figure 2 list the proposed process of generating protein key.

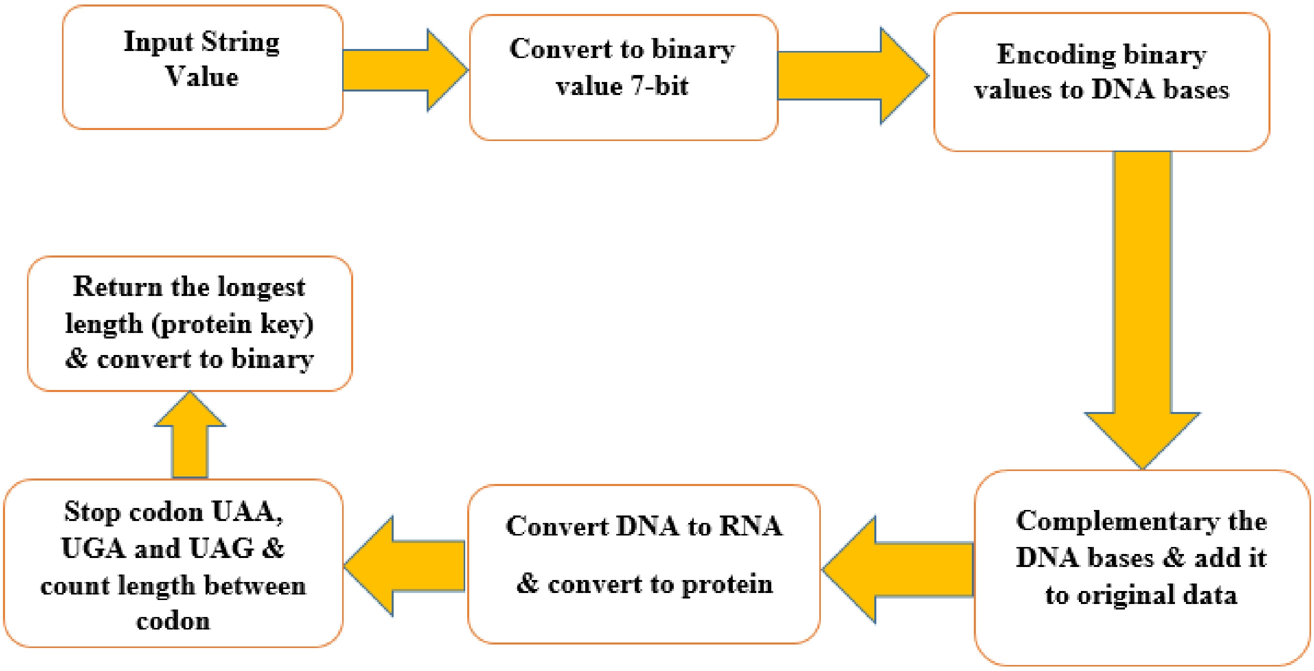

Input String Value
Convert to binary value 7-bit
Encoding binary values to DNA bases
Complementary the DNA bases & add it to original data
Convert DNA to RNA & convert to protein
Stop codon UAA, UGA and UAG & count length between codon
Return the longest length (protein key) & convert to binary

Figure 2: Proposed Process For Generating A Protein Key

Table 1 Nucleotide Coding

Bit-1	Bit-2	DNA Code
0	0	A
0	1	G
1	0	C
1	1	T

Table 2 Complement Of Nucleotide

DNA	Complementary
A	T
G	C
C	G
T	A

Having a protein key makes the encryption and decryption process more secure and easy to do. In the encryption process, the AES algorithm is used with some modifications. One of the most disadvantage of AES that due to their intrinsic features, images have poor encryption effects like high data redundancy and bulky data capacity. To overcome this problem modification of AES with DNA computing has been proposed. Firstly, the input image with size m*n pixels will be separated into a 4*4 matrix. Secondly, each matrix is encrypted with the AES algorithm. The process of encryption is made up of two rounds: first round, round reduced of the AES 128-bit block cipher used. Second round, protein key generated in previous steps used in this process. The protein key XORed with the output of the first block and so on for all blocks (Algorithm 1 describes pseudo-code for the proposed Encryption method.

```
Input Plain image n×m size, Protein key
Output Encrypted image m×n
For each 4×4 pixel in image
    AddRoundKey(pixel,proteinkey )
    For each bit in key generation do
        SubBytes(pixel)
        ShiftRows(pixel)
        MixCoulmns(pixel)
        AddRoundKey(pixel,key generation)
    End loop
    SUBBytes(pixel)
    ShiftRows(pixel)
    AddRoundKey(pixel,key generation)
End loop
XOR(pixel, DNA key)
```

Algorithm 1 Pseudo-Code For The Encryption Algorithm

5. Proposed Algorithm Performance Analysis

MATLAB is used to implement the proposed method under Windows 10 environment with core 3 and 4 Gb memory. CT chest for infected people has been selected for testing the proposed method. Figure 3 list the encryption images for the selected plain images. The encryption process uses the same protein key as the decryption process. The encrypted images prove that the proposed method able to achieve a high level of security and also indicate that the encrypted image will resist all kinds of known attacks as its encryption quality doesn’t depend on the protein key and plain image. To illustrate the encryption method quality, the histogram analysis of cipher images will be used. This property makes statistical attacks difficult in images. Figure 4 illustrate a uniformly distributed histogram for cipher images. After the encryption process, the histogram was very flatted for the ciphertext image. In other words, the ciphertext images are randomly like. This proved that the encrypting method is a perfect image encryption method, and the intruder is unable to identify the plain images.

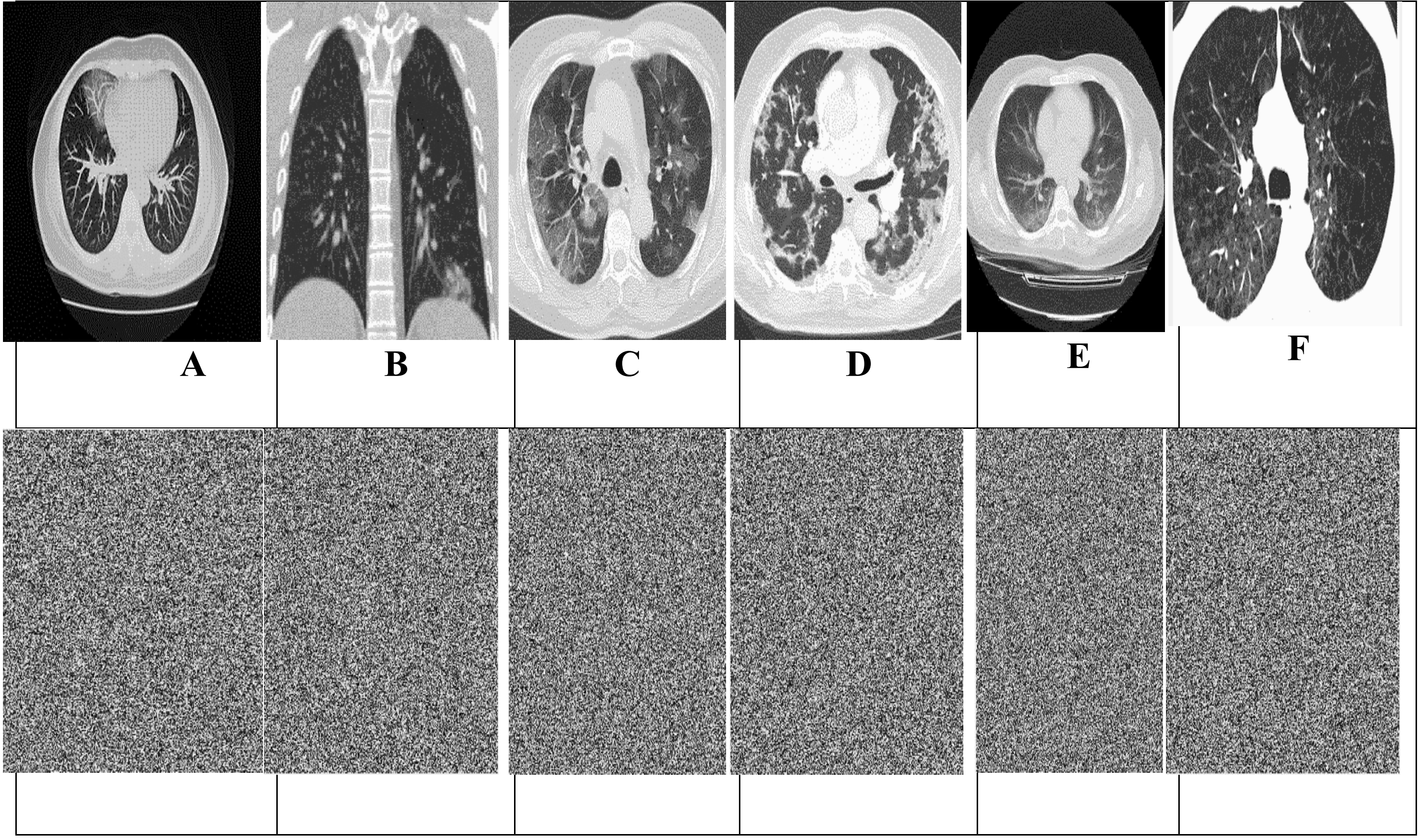


Figure 3 Plain Images And Cipher Images

6. Evaluation And Experimental Results

The proposed algorithm is evaluated based on four parameters: -

1. Number of Pixel Change Rate (NPCR)

Estimate the degree of sensitivity for the encryption method with changing of the pixels in the original image. If the NPCR value is close to 100%, this indicates that the proposed method can resist the statistical attacks and have a high level of security. NPCR can be evaluated from this equation: -

$$\mathrm{NPCR} = \frac{\sum_{i=0}^{n} \quad \sum_{i=0}^{m} D(i,j)}{n \times m} * 100 \qquad \text{Eq. (1)}$$

Where D(i,j) is the dimension of the image at positions i and j and n is the width of image and m is the height of the image. The NPCR for the proposed method is discussed in table 4. The results indicated that the proposed method achieves more sensitivity since the NPCR value above 99.5%.

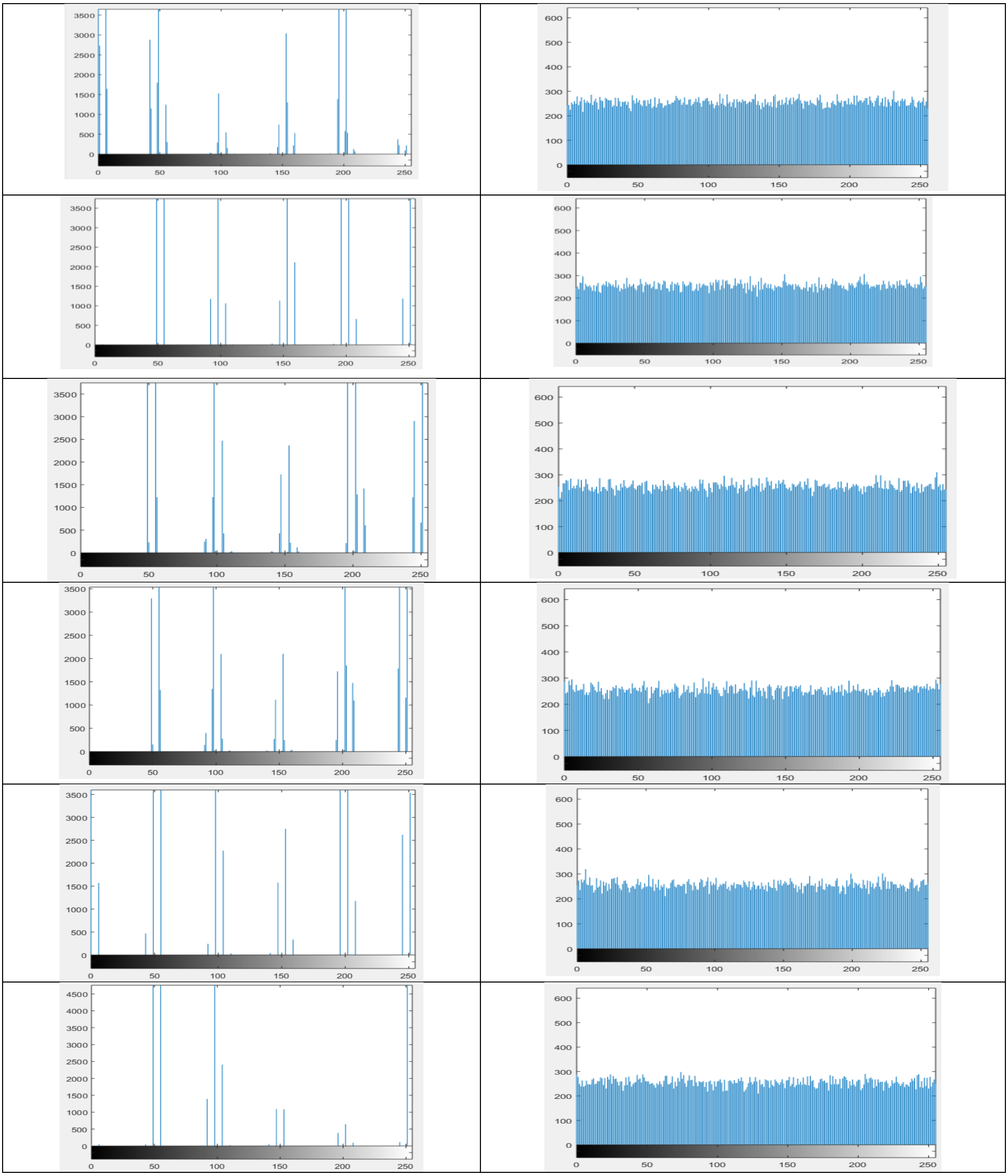


Figure 4: Histogram For Plain Images And Cipher Images

2. *Correlation coefficient*

The low correlation coefficient between encrypted image pixels, the high security the algorithm achieved. The correlation coefficient can resist statistical attacks when its value is closed to zero. For each adjacent pairs, the Correlation coefficient can be calculated as follows:

$$P = \frac{\sum_{i=0}^{n} \sum_{j=0}^{m} (X(i,j) - E(X))(Y(i,j) - E(Y))}{\sqrt{\sum_{i=0}^{n} \sum_{j=0}^{m} (X(i,j) - E(X))2 \sum_{i=0}^{n} \sum_{j=0}^{m} (Y(i,j) - E(Y))2}} \qquad \text{Eq. (2)}$$

where X(i,j) is the row and column pixels in the plain image, Y(i , j) is the row and column pixels in the cipher image. E(X) is the average of the plain image, E(Y) is the average of the cipher image. Table 3 lists the Correlation Coefficient for both AES only and AES + Protein key the results indicate that AES with Protein key decreases the correlation value which makes the proposed method achieve a high level of security.

Table 3. Correlation Coefficient

File Name	AES only	AES+Protein
CT Chest1	0.0084	0.0034
CT Chest2	0.0019	-0.0070
CT Chest3	0.0079	0.0019
CT Chest4	0.0031	0.0035
CT Chest5	0.0067	0.0033
CT Chest6	-0.0047	-0.0018

3. *Unified Average Changing Intensity (UACI)*

The greater the UACI value proposed method score, the highest level of security and avert a differential attack from an intruder attain. UACI can be calculated from the following equation: -

$$UACI = \frac{1}{n*m} * \frac{\sum_{i=0}^{n} \sum_{j=0}^{m} |C1(i,j) - C2(i,j)|}{256} * 100 \qquad \text{Eq. (3)}$$

where C1, C2 are two images with the same size if they are not the same then the value of D(i,j) equal 1 if not then D(i,j) equals 0. Table 4 list the UACI for the proposed method. The values above 30\% so the method can apply a high level of security. D(I,j)=|c1,c2|.

4. *Entropy*

This parameter estimates the amount of information lost in the decryption process. Many researchers get the entropy value close to 8 as discussed here [14-21]. When the value close to 8 this means that there is a very little amount of information that has been lost. The entropy value can be calculated for M messages as follows:

$$H(m) = \sum_{i=1}^{n} p(mi) log2 \frac{1}{p(mi)} \qquad \text{Eq. (4)}$$

Where p (m) in the image is the probability of pixel value. The Entropy value for the Plain image and encrypted image discussed in table 5. With regards to results, cipher images have an entropy value near 8 this indicates that the proposed method applies a high level of sensitivity and security.

Table 4. NPCR and UACI

File Name	NPCR	UACI
CT Chest1	99.5895	39.5097
CT Chest2	99.6445	32.7303
CT Chest3	99.5605	33.7530
CT Chest4	99.5834	36.2310
CT Chest5	99.5621	36.3790
CT Chest6	99.6063	36.6857

Table 5. Entropy

File Name	Plain	Cipher
CT Chest1	3.7660	7.9968
CT Chest2	3.1894	7.9969
CT Chest3	3.6335	7.9972
CT Chest4	3.6588	7.9974
CT Chest5	3.3835	7.9970
CT Chest6	2.5372	7.9972

7. CONCLUSION

This study proposes a key generation from protein codon. This key is used for each encryption and decryption process. It also uses AES with two rounds for these processes. The histogram is used to estimate the quality of the proposed method. The histogram of cipher images illustrates a uniformly distributed histogram. The cipher image histograms become very flat after encryption. This indicates that the proposed method can efficiently encrypt images and the intruder is unable to identify the plain images. To estimate the sensitivity of the proposed algorithm, four parameters used NPCR, UACI, Entropy, and Correlation Coefficient. The results approve that the proposed method can improve the security level and sensitivity as all cipher images have above 99.5% NPCR and less correlation near to zero and also entropy near to 8 and UACI above 30%.

REFERENCES

1. A. Kahate, Cryptography, and Network Security, Third Edition, Mc Graw Hill, 2016.
2. Z. Cao, L. Liu, Z. Guo, Ruminations on attribute-based encryption, International Journal of Electronics and Information Engineering, vol. 8, no. 1, pp. 9–19, 2018.
3. T. Gulom, The encryption algorithms GOST28147- 89-IDEA8-4 and GOST28147-89-RFWKIDEA8-4, International Journal of Electronics and Information Engineering, vol. 6, no. 2, pp. 59-71, 2017.

4. T. Gulom, The encryption algorithms GOST28147- 89-IDEA8-4 and GOST28147-89-RFWKIDEA8-4, International Journal of Electronics and Information Engineering, vol. 8, no. 1, pp. 20–31, 2018.
5. T. Gulom, the encryption algorithm AES-PES16-1 and AES-RFWKPES16-1 based on network PES16-1 and RFWKPES16-1, International Journal of Electronics and Information Engineering, vol. 3, no. 2, pp. 53–66, 2015.
6. A. Mermaid, T. Gulom, the encryption algorithm AES-RFWKIDEA32-1 based on network RFWKIDEA32-1, International Journal of Electronics and Information Engineering, vol. 4, no. 1, pp. 1–11, 2016.
7. Bruce Schneier; John Kelsey; Doug Whiting; David Wagner; Chris Hall; Niels Ferguson; Tadayoshi Kohno; et al., The Twofish Team's Final Comments on AES Selection (PDF). Archived (PDF) from the original on 2010-01-02,2002.
8. L. M. Adleman, Molecular computation of solutions to combinatorial problems, Science, vol. 266, no. 5187, pp. 1021-1025, 1994.
9. T. Mahalaxmi, B. B. Raj, J. F. Vijay, Secure data transfer through DNA cryptography using a symmetric algorithm, International Journal of Computer Applications, vol. 133, no. 2, pp. 19-23, 2016.
10. N. Gulati, S. Kalyani, Pseudo DNA cryptography technique using OTP key for secure data transfer, International Journal of Engineering Science and Computing, vol. 6, no. 5, pp. 5657-5663, 2016.
11. T. Purusothaman, K. Saravanan, DNA-based secret sharing algorithm for a multicast group, Asian Journal of Information Technology, vol. 15, no. 15, pp. 2699-2701, 2016.
12. E. S. Babu, M. H. M. K. Prasad, C. N. Raju, Inspired pseudo biotic DNA based cryptographic mechanism against adaptive cryptographic attacks, International Journal of Network Security, vol. 18, no. 2, pp. 291-303, 2016.
13. K. Chiranjeevi, S. L. Kumar, R. Paspula, Hidden data transmission with variable DNA technology, International Journal of Electronics and Information Engineering, vol. 7, pp. 41-52, 2017.
14. Arroyo, David, Rhouma Rhouma, Gonzalo Alvarez, Veronica Fernandez, and Safia Belghith, On the skew tent map as the base of a new image chaos-based encryption scheme, In Second Workshop on Mathematical Cryptology, pp. 113-117, 2008.
15. Yin, Ruming, Jian Yuan, Qiuhua Yang, Xiuming Shan, and Xiqin Wang, Gemstone: a new stream cipher using coupled map lattice, In Information Security and Cryptology, pp. 198-214, 2010.
16. Öztürk, İsmet, and İbrahim Soğukpınar, Analysis and comparison of image encryption algorithms, International Journal of Information Technology 1, no. 2, pp. 108-111,2004.
17. Jawad, Lahieb Mohammed, and Ghazali Bin Sulong, A review of color image encryption techniques, International Journal of Computer Science Issues 10, no. 6, pp. 266-275, 2013.
18. Mulualem, Getachew my habit, Compression, and encryption for satellite images: a comparison between squeeze cipher and spatial simulations, (2015).
19. Johnson, Mark, Prakash Ishwar, Vinod Prabhakaran, Daniel Schonberg, and Kannan Ramchandran, on compressing encrypted data, Signal Processing, IEEE Transactions on 52, no. 10, pp.2992-3006, 2004.
20. Kline, Demijan, Carmit Hazay, Ashish Jagmohan, Hugo Krawczyk, and Tal Rabin, On compression of data encrypted with block ciphers, Information Theory, IEEE Transactions on 58, no. 11, pp.6989-7001,2012.
21. Zhou, Jiantao, et al, Designing an efficient image encryption-the compression system via prediction error clustering and random permutation, Information Forensics and Security, IEEE Transactions on 9.1, ppt.39-50,2014.
22. Pramod Pavithrana ,Sheena Mathewa, Suyel Namasudrab, Pascal Lorenzc, A novel cryptosystem based on DNA cryptography and randomly generated mealy machine , Elsevier, Volume 104, May 2021.