Color Image Steganography Based on Artificial DNA Computing

Najat H. Qasim¹, Sahera A. Sead²
E-mail: najat.qasim@uobasrah.edu.iq, sahera.sead@uobasrah.edu.iq
¹Department of English language, College of Arts, University of Basrah, Iraq.
²Department of information system, College of Computer sciences and Information Technology, University of Basrah, Iraq

Keywords: steganography, DNA, image, LSB, encoding, security

Received: March 30, 2023

Modern genetic engineering developments have made it possible for artificial DNA strands to be included in living cells of creatures. Many methods of artificial insertion have been developed using DNA, which has excellent data storage capacity. Most of these techniques are used to encode text data, while there has been little research on encoding other types of media. Methods for encoding images have very little studied, and most of them are dedicated to black-and-white images. The proposed method focuses on encoding a secret color image and then embedding it in another color image, this comprises two levels of security. The first level is provided by converting binary color images into DNA sequences. A second level is provided by embedding the bits of DNA sequence into LSBs of the cover image to generate the stego image. Extraction process is in the reverse procedure. The proposed method is significantly efficient, according to the experimental results.

1 Introduction

The emergence of the big data era has resulted in the everyday generation of various digital images that are highly informational. As a result, the security concerns with digital images have gotten much worse. However, extensive security flaws, in addition to, high correlation between pixel points and their large data capacity then, traditional data encryption methods are not appropriate for image encryption. As a result, scientists started looking for novel approaches to image encryption. Cause to DNA’s special properties, great parallelism, and high information density, DNA coding is a popular topic in the images encryption field. The use of DNA computing in cryptography and steganography has been identified as a potential technology that could provide a new hope for unbreakable algorithms. The combination of encryption and steganography can provide a high level of security and protect the privacy of information. Steganography is the science of concealing data in a cover, which can be audio, text, image, or video.

In our work, we present a new method that combines DNA encryption and LSBs steganography. Where, the color image encrypts and compressed before hiding them in the cover image.

The rest of this research arranges as: Section 2 present the related works. Section 3 explains the DNA encoding rule. We explain the suggested algorithm in section 4. We describe research results and analyses in section 5. The discussion is in section 6. We present the conclusions of this work in section 7. The future work introduce in section 8.

2 Related works

Researchers have improved DNA encryption techniques to increase the performance of cryptography algorithms. They apply DNA encryption in many articles on data encryption. While compared with a conventional system, DNA is perfect for developing an efficient and secure cryptosystem due to its enormous information storage capacity, powerful parallel processing capability, and low energy usage. For realizing image coding, current studies use DNA encryption methods, a DNA encryption procedure (addition, subtraction, complement, XOR, etc.), or a combination of several coding operations. In other words, no existing study explains why a specific DNA encryption method (static or dynamic), the DNA encryption process, or a combination of several encryption methods was chosen to achieve image encryption.

Today, scientists are very interested in experimenting with DNA to improve encryption techniques and offer a more efficient method for encrypting color images.

3 DNA encoding rule

Adleman originally applied DNA computing in the field of encryption in 1994, ushering in a new stage for data processing. DNA encryption is a novel field that is currently at the forefront of global research in cryptography. DNA molecules are highly energy-efficient, have a high storage density, and can run in enormous parallel. As a result, DNA-based image encryption techniques have certain features that other
cryptography techniques do not have. The technique of mapping the nucleotide sequence that makes up a DNA strand is known as DNA sequencing. Adenine (A), thymine (T), guanine (G), and cytosine (C) are the four nucleic acid bases that combine to create DNA sequence components and the building blocks of genetic coding. A DNA sequence has the form of a binary string, and every two nucleic acid bases on either side of the string are complementary, according to the criteria that A and T are complementary to each other as well as G and C. Only eight possible DNA coding combinations are suitable for the complementarity principle. Table 2 includes a list of these. Where, the binary numbers 00, 01, 10, and 11 are complementary and represent the four deoxynucleotides A, C, G, and T respectively. We are percipient that every pixel in a digital image can be represented by an 8-Bit binary value. Then each pixel in a digital image can be represented by a string of nucleotides. So, sequence of nucleotides corresponds to the binary number (11100100) according to the preceding principles (TGCA).

Table 1: Represent the state-of-the-art about DNA encryption for color images

<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>Method</th>
<th>Suggested work</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>[1]</td>
<td>DNA encryption algorithm based on DNA sequence and multiple improved 3D shader</td>
<td>Suggested DNA encryption algorithm based on DNA sequence and multiple improved 3D shader</td>
</tr>
<tr>
<td>2018</td>
<td>[2]</td>
<td>Color image encryption scheme using CRL and DNA sequence operations</td>
<td>Presented a color image encryption scheme using CRL and DNA sequence. The pixel matrix was created by CRL, and the color image was encoded using DNA sequence operations.</td>
</tr>
<tr>
<td>2017</td>
<td>[3]</td>
<td>DNA-based color image encryption algorithm using DNA sequence operations</td>
<td>Developed a special encryption method based on DNA sequence operations. The initial color image was split into three frames (R, G, and B), and these frames were subsequently transformed into DNA matrices.</td>
</tr>
<tr>
<td>2018</td>
<td>[4]</td>
<td>DNA-based color image encryption algorithm using DNA sequence operations</td>
<td>Presented a color image encryption algorithm based on DNA sequence operations. The pixel matrix was created by a special encryption method that utilized DNA sequence operations. The initial color image was split into three frames (R, G, and B), and these frames were subsequently transformed into DNA matrices.</td>
</tr>
<tr>
<td>2019</td>
<td>[5]</td>
<td>DNA-based color image encryption algorithm using DNA sequence operations</td>
<td>Presented a DNA-based color image encryption algorithm. The pixel matrix was created by a special encryption method that utilized DNA sequence operations. The initial color image was split into three frames (R, G, and B), and these frames were subsequently transformed into DNA matrices.</td>
</tr>
<tr>
<td>2020</td>
<td>[6]</td>
<td>DNA-based color image encryption algorithm using DNA sequence operations</td>
<td>Presented a DNA-based color image encryption algorithm. The pixel matrix was created by a special encryption method that utilized DNA sequence operations. The initial color image was split into three frames (R, G, and B), and these frames were subsequently transformed into DNA matrices.</td>
</tr>
</tbody>
</table>

Table 2: DNA coding Rule.

<table>
<thead>
<tr>
<th>Rule</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>01</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>10</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>11</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>G</td>
</tr>
</tbody>
</table>

4 Suggested method

Steganography is the science of conveying a message concealed inside a cover medium. So that only the sender and the intended recipient are aware of its presence. Steganography strives to hide the existence of a message, while cryptography only seeks to render a message unreadable. The suggested method consists of two levels of security: the first level is the encryption stage, which constructs DNA in the encryption mechanisms, and the second is the hiding stage. The following is an explanation of each level:

4.1 The DNA encryption algorithm

At this stage first, we secure the color image using the DNA coding rule. Second, we compress the binary image and then hide it in the least significant bit of the host image. The encryption algorithm summarizes in the following steps:

1. Enter colored secret image.
2. Specifies the secret image size to 256×256.
3. Partition the secret image into three layers (Red, Green, and Blue).
4. Using the DNA method to encode the secret image as follows:
   A. Convert the secret image data into binaries while working on the first (red) layer.
   B. In the first layer, convert four LSB of each byte to the decimal number d and then calculate the remainder as R = mod(d, 8). Example: R= mod(15,8)=7.
   C. The output of the prior steps ranges from 0 to 7, representing the base number used for encryption, according to Table 2. Example: 01111111= AGGG (using Rule 7).
   D. Every 2 bits in the current byte replace with its corresponding symbol in Table 2.
   E. Following the repetition of steps (B, C, and D) on all the red layer’s bytes, we repeat step 4 on each of the green and blue layers.
   F. So that we can acquire RDNA, GDNA, and BDNA, respectively, and their sizes are (4, N, M).
5. Converting DNA codes into binary based on the Davis method, we give ascending order to the bases by Davis. These results are C = 1, T = 2, A = 3, and G = 4. This approach reduces the binary digits of the bit-mapped image into fewer DNA base codes, with each base indicating how many times each code is repeated. This method is widely used in data compression.

The following procedures are used to produce the symbols of the coding table. Which are formed upon implementation and have a separate coding table for each secret image.
- Analyze the output of the preceding coding step to determine how many times each DNA symbol is repeated. We use the following illustration to explain this process. Table 3 will be the result of this procedure.

- The representations of Table 3 arrange in descending order, so the results presented in Table 4 as:

- The four symbols ((T, C, A, and G)) are each assigned a code since the most common symbol uses the fewest binaries. Where fewer binaries will be required to reduce the DNA codes. As we see in Table 5 and Table 9.

<table>
<thead>
<tr>
<th>Table 3: Analyze the output of DNA Encoding Representations</th>
<th>Base</th>
</tr>
</thead>
</table>
| 50 | T  
| 30 | C  
| 20 | A  
| 40 | G |

<table>
<thead>
<tr>
<th>Table 4: Descending order of Table 3 Representations</th>
<th>Base</th>
</tr>
</thead>
</table>
| 50 | T  
| 40 | G  
| 30 | C  
| 20 | A |

<table>
<thead>
<tr>
<th>Table 5: DNA Encoding scheme based Davis Bit sequence</th>
<th>Base</th>
</tr>
</thead>
</table>
| 1 or 0 | T  
| 11 or 00 | G  
| 111 or 000 | C  
| 1111 or 0000 | A |

Then, the current DNA code will replace with a few zeros or ones. Depending on Table 5, as well as depending on the binary value of the symbol that accepted it. For example, TGATC= 1001110111.

4.2 The steganography algorithm
This level applies the least significant bits (LSB) technique to hiding the binary codes of a secret color image in the LSB of a host image. The hiding algorithm summarizes in the following steps:
1. Select the colored host image as input.
2. Specify the host image size to 300x300.
3. Partition the host image into three layers (Red, Green, and Blue).
4. Transform the host image data into binaries and work on the first (red) layer.
5. Load sequentially bits from the first layer of the encrypted image into the third least significant bit of the host image. Where the first and second LSB of the stego image will be changing as follows:

- Case 1: If the third LSB of the host image is zero and the current bit of the encrypted image is one. So, the values of the (1st and 2nd) least significant bits of the stego image will become zeros, as demonstrated the example in figure 1.(a).

- Case 2: If the third LSB of the host image is one and the current bit of the encrypted image is zero. So, the values of the (1st and 2nd) least significant bits of the stego image will become ones, as demonstrated the example in the figure 1.(b).

- Case 3: In the case of the absence of either of the two mentioned conditions, the current bit of the encrypted image will only include the third least significant bit of the stego image without any other changing.

6. Apply steps (4 and 5) over two others layers (green and blue).
7. Obtain the stego image.

4.3 The recovered algorithm
This algorithm represents the reverse of the second level of security where the stego image is used to generate the encrypted image.

4.4 The decryption algorithm
We typically use a decryption algorithm, which is the opposite of an encryption algorithm. This algorithm is the inverse of the first level of security. The encrypted image and the secret keys are required to restore the original image.
5 Experimental results and analysis

We carried many experiments results to present the performance and validity of the suggested method in two levels of security. The first test; starts with the results that evaluate the encryption/decryption processes on the different images. We test three cover images (Lena, Baboon, and Pepper) to reveal the power of DNA encryption and the hiding algorithm. To verify the security and performance of the algorithm, we must test and evaluate this algorithm based on the properties of the stego image. A good algorithm would produce a stego image that meets the evaluation metrics requirements. We can classify metrics into two categories. The first group evaluates the efficiency of the substitution process, which includes histograms, entropy coefficients, and correlation. The second group evaluates the approach's ability to propagate the original image. That includes MSE, PSNR, NPCR, and UACI. The suggested method is compared to the performance of competing methods based on a variety of parameters. The results showed that the suggested method is superior to other methods in terms of safety. Correlation comparison results (color images calculated by averaging the values of the horizontal, vertical, diagonal, red, green, and blue components) and information entropy comparison these results are presented in Table 7.

5.1 Mean Square Error (MSE) Analysis

The amount of distortion in the image that indicates the variance between the cover and a stego image is measured using the MSE. The quality of the image is good if the MSE value is low.

\[ MSE = \frac{[I_1(M,N) - I_2(M,N)]^2}{M \times N} \]  

(1)

\(I_1(M, N)\) denotes the host image pixel in the (M, N) point, where M and N define the rows and columns numbers in the host image, and \(I_2(M, N)\) denotes the stego image pixel in the (M, N) location. Table 6 presents results obtained from the suggested method; as can be seen, the MSE values are low, indicating very acceptable stego image quality.

5.2 Peak Signal to noise ratio (PSNR) analysis

The Peak Signal to Noise Ratio is a metric used to calculate the decibel level of imperceptibility. It compares the original host images' and stego images' quality. Refers to a tiny difference between a host and stego images when the PSNR value is high. The algorithms used in steganography are designed to boost PSNR. According to the following equation, PSNR is:

\[ PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \]  

(2)

The results of the experiments and the analysis discussed above show that the proposed method yields the highest PSNR values, making it hard to discover the hidden image. The results in Table 6 show the algorithm offers higher protection against threats and improved security.

Table 6: Illustrates the experimental results of PSNR and MSE

<table>
<thead>
<tr>
<th>Cover Image</th>
<th>Cover Image Size</th>
<th>MSE</th>
<th>PSNR</th>
<th>Secret Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>463x516</td>
<td>0.000977</td>
<td>88.2316</td>
<td><img src="image" alt="Lena" /></td>
</tr>
<tr>
<td>Baboon</td>
<td>463x516</td>
<td>0.000998</td>
<td>88.1988</td>
<td><img src="image" alt="Baboon" /></td>
</tr>
<tr>
<td>Pepper</td>
<td>256x256</td>
<td>0.000975</td>
<td>88.2406</td>
<td><img src="image" alt="Pepper" /></td>
</tr>
</tbody>
</table>

5.3 Information entropy analysis

We measure the ambiguity of a system using information entropy. That quantifies the unpredictability of an image plus the quantity of information included within it. A distribution of pixel values in an image can be determined using information entropy. The greater the
importance of using encryption, the better the results. So, entropy $H(d)$ for data $d$ will calculate as follows.

$$H(d) = \sum_{i=1}^{l} p(d_i) \log_2 \left( \frac{1}{p(d_i)} \right)$$  \hspace{1cm} (3)

where $p(d_i)$ represents the probability of $d$.

### 5.4 Correlation analysis

Correlation analysis calculates the degree of correspondence between the host image before and after hiding. When the correlation coefficients of the pixels in the host image are the least possible, the color image-hiding technique will withstand statistical attacks. Equations (4), (5), and (6) determine the correlation coefficients between two neighboring pixels along their diagonal, horizontal, and vertical axes:

$$corr_{xy} = \frac{\text{cov}(x,y)}{\sqrt{\sigma_x \sigma_y}}$$  \hspace{1cm} (4)

$$\text{cov}(x, y) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})$$ \hspace{1cm} (5)

$$\sigma_x = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2, \quad \sigma_y = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2$$ \hspace{1cm} (6)

Correlation coefficients are calculated for several rates of close pairs (horizontal, vertical, and diagonal) chosen randomly from the host image. Table 7 shows the comparative results of the suggested method for the (Entropy, Correlation) between the host and stego images. Correlation studies show this technique has excellent steganography.

### 6 Discussion

Many of the previous studies used hybrid approaches based on DNA to increase image encryption and resistance to attacks, as illustrated in Table 1.

But, when transmitting data over an unsafe public connection, steganography is frequently preferred over encryption. We cannot achieve security via cryptography only. So, several computing-related settings are suggested for using various DNA sequence biological properties in steganography and secure cryptography. Researchers have developed DNA-based methods for hiding data that offer unmatched confidentiality and protection without limiting adaptability or storage capacity. And they have created unique data-hiding strategies based on DNA after learning about the biological characteristics of DNA sequences. Such has prompted the creation of a whole new scientific area.

To ensure secure digital color images, including those used in medicine, satellite remote sensing, social networking, etc., in this research, we suggested a new encoding algorithm to generate a DNA code for a confidential image. Furthermore, we produced an encoding technique for reducing DNA data storage, which is a relatively recent advancement in the world of digital data storage. Table 9 shows the obtained results. Then, we proposed a new approach to hide the bits of DNA sequence in the least significant bits (LSB) of the cover image.

### Table 7: Experimental results of the suggested method for (Entropy, Correlation) between the host image and stego-image

<table>
<thead>
<tr>
<th>Cover Image</th>
<th>Cover Image Size</th>
<th>Entropy Cover Image</th>
<th>Entropy stego Image</th>
<th>Correlation Cover Image</th>
<th>Correlation stego Image</th>
<th>Secret Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>463×516</td>
<td>7.2625</td>
<td>7.1630</td>
<td>0.9626</td>
<td>0.9644</td>
<td>Lena</td>
</tr>
<tr>
<td>Baboon</td>
<td>463×516</td>
<td>7.3864</td>
<td>7.2761</td>
<td>0.9666</td>
<td>0.9670</td>
<td>Baboon</td>
</tr>
<tr>
<td>Peppers</td>
<td>256×256</td>
<td>7.6586</td>
<td>7.5402</td>
<td>0.8009</td>
<td>0.7973</td>
<td>Peppers</td>
</tr>
</tbody>
</table>

### 7 Conclusions

In this paper, we suggested a color image steganography constructed using artificial DNA computing to provide more security to the system. Maintaining a high level of robustness against attacks is the objective of the suggested technology. Differential and statistical analyses we used with various tools demonstrate the security of our algorithm. To test used different image formats and a variety of image sizes to compute the (PSNR) and (MSE). The results and analysis proved that the proposed algorithm for differential analysis is resilient against attacks. Table 8 shows the obtained results compared to the other techniques. The results of the tests on the three standard images where a higher PSNR value between the host
images and the stego images and a lower MSE value, which shows that the algorithm’s security is robust and this technique has excellent steganography.

Table 8: PSNR and MSE comparison results based on three colored images (Lena, Baboon, and Peppers).

<table>
<thead>
<tr>
<th>Images</th>
<th>Method</th>
<th>PSNR</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>Ours</td>
<td>88.2316</td>
<td>0.0000977</td>
</tr>
<tr>
<td></td>
<td>[44]</td>
<td>83.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[25]</td>
<td>65.159</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>[21]</td>
<td>50.12</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>[10]</td>
<td>57.49</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>[5]</td>
<td>60.4429</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>[46]</td>
<td>82.59</td>
<td>0.173</td>
</tr>
<tr>
<td>Baboon</td>
<td>Ours</td>
<td>88.2306</td>
<td>0.0000973</td>
</tr>
<tr>
<td></td>
<td>[44]</td>
<td>83.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[25]</td>
<td>66.115</td>
<td>0.00089</td>
</tr>
<tr>
<td></td>
<td>[21]</td>
<td>46.01</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>[15]</td>
<td>45.9012</td>
<td>1.5887</td>
</tr>
<tr>
<td></td>
<td>[10]</td>
<td>58.19</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>[5]</td>
<td>60.4534</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>[46]</td>
<td>82.59</td>
<td>0.173</td>
</tr>
<tr>
<td>Peppers</td>
<td>Ours</td>
<td>88.1981</td>
<td>0.0000998</td>
</tr>
<tr>
<td></td>
<td>[44]</td>
<td>83.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[25]</td>
<td>68.0086</td>
<td>0.0088</td>
</tr>
<tr>
<td></td>
<td>[21]</td>
<td>50.00</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>[15]</td>
<td>45.0216</td>
<td>2.0621</td>
</tr>
<tr>
<td></td>
<td>[10]</td>
<td>52.48</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>[5]</td>
<td>60.5005</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>[45]</td>
<td>71.7357</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td>[46]</td>
<td>82.59</td>
<td>0.153</td>
</tr>
</tbody>
</table>

Table 9: Results of the encryption algorithm that reduced DNA data storage.

<table>
<thead>
<tr>
<th>Secret Images</th>
<th>Images Sizes</th>
<th>Number of Bits in Binary Secret Images</th>
<th>Number of DNA Codes in Secret Images</th>
<th>Number of Bits in Encrypted Secret Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>75x75</td>
<td>135000</td>
<td>5625</td>
<td>54666</td>
<td></td>
</tr>
<tr>
<td>75x75</td>
<td>135000</td>
<td>5625</td>
<td>43437</td>
<td></td>
</tr>
<tr>
<td>75x75</td>
<td>135000</td>
<td>5625</td>
<td>47106</td>
<td></td>
</tr>
<tr>
<td>75x75</td>
<td>135000</td>
<td>5625</td>
<td>47704</td>
<td></td>
</tr>
</tbody>
</table>

8 Future work
We can expand this work in the future to include:
Both speech and video can be encoded using the method.
To secure the image and maintain the proper size, it is advised to use image encryption and image data reduction technologies.

References
https://doi.org/10.1109/access.2020.2991420

https://doi.org/10.1107/s11042-017-4534-z

https://doi.org/10.1107/s11042-018-6989-y

https://doi.org/10.1016/j.optlaseng.2019.105851

https://doi.org/10.1107/s11042-018-6758-y

https://doi.org/10.1016/j.ijleo.2018.01.064

https://doi.org/10.3390/e22020158

https://doi.org/10.1109/access.2020.2970981

https://doi.org/10.1088/1757-899x/518/5/052003

https://doi.org/10.1007/s00500-021-06459-w


