IMAGE ANONYMIZATION USING CHAOTIC MAPPING AND DNA ENCODING

Course Title: Data privacy

Slot: A2

Submitted by

O Nikhil Kumar J - 18BCl0091

O Mugesh – 20BCE0300

O Archit Harsh - 18BCI0153

ABSTRACT

Images are a 2D array of pixels whose value is stored in 24 bits containing RGB values. In order to obtain an obscured image which still has utility we use logistic function mapping to scramble the pixels. The range of scrambling in an m x n image can be confined into p x q blocks. The pixels are then transformed into DNA format and complemented with a key which is known as DNA encoding. The resulting image obscures a lot of the original details but still has utility based on the requirements. The algorithm can be implemented in python without special libraries using in-built algorithms.

1. INTRODUCTION

With the advancements in technology, images are an important source of data. Simple nuances in images can be recognised and utilised by softwares to obtain valuable information. Big data companies predict that image recognition and analysis will become increasingly important in areas such as health care, environmental monitoring, human/object recognition and satellite images. With rising importance of images in data analytics, companies have incentive to use their image data for commercial purposes.

We propose a method to alter the image in such a way that its identifying characteristics are scrambled while retaining some of the properties that can be crucial for data analysis. We confuse pixel location in a particular range which obscures the image using the logistic function which is highly random and particularly sensitive to initial values, hence descrambling is unworkable.

However, although chaotic encryption technology has many advantages, it also has a number of deficiencies. For example, at present, most chaotic encryption algorithms confuse the single image pixel value or location, but the utilization of only one of the two strategies do not ensure high security for the image, and thus it is easy for attackers to crack an encrypted image by simply using the pixel comparison method. Hence we also combine DNA encoding which possesses unique advantages that the traditional cryptographic algorithms do not have. The final output is an scrambled and encrypted image that can be utilized without identification.

2. RELATED WORK

After the survey of various methods used for image encryption, we came through a few of these like Image encryption using chaotic maps, DES...

1. Zhang, DongXin Fang, Honge Ren, ''Image Encryption Algorithm Based on DNA Encoding and Chaotic

Maps": This paper discusses image encryption algorithm based on DNA sequences combined with chaotic maps. For any size of the original grayscale image, the rows and columns are fist exchanged by the arrays generated by a logistic chaotic map. Secondly, each pixel that has been confused is encoded into four nucleotides according to the DNA coding. Thirdly, each nucleotide is transformed into the corresponding base pair a random number of time(s) by a series of iterative computations based on Chebyshev's chaotic map. They show that the seecurity analysis not only has good encryption effect, but also has the ability to repel exhaustive, statistical, differential, and noise attacks.

2. A New Chaotic System for Image Encryption (2012): In this script, Long Bao and Yicong Zhou have suggested a new chaotic system that constitutes the three distinct one-dimensional chaotic

maps. The suggested technique applies the Logistic map as a controller to choose the Tent map or a Sine map to generate random sequences. Thereafter, the imparted algorithm utilizes the substitution permutation network (SPN) structure to obtain the confusion and diffusion property. This scheme uses 240 bit key for large key space. Mainly, this key contains all parameter settings and the initial values of the new chaotic system, and excessive sensitivity in key changes for encryption and decryption. Consequently, the proposed approach provides an excellent security against the brute force attack as well as extreme key sensitivity and chaotic behaviour.

3. Image Encryption Using Block-Based Transformation Algorithm: Here a block-based transformation algorithm based on the combination of image transformation and a well-known encryption and decryption algorithm called Blowfish is used. The original image was divided into blocks, which were rearranged into a transformed image using a transformation algorithm presented here, and then the transformed image was encrypted using the Blowfish algorithm. The results showed that the correlation between image elements was significantly decreased by using the proposed technique.

3. PROPOSED WORK

Chaotic encryption method is utilized to provide a pixelated image. We propose rearrangement of pixels by using logistic mapping function which is used to create pseudo random sequence based on input value. The range of rearrangement of the pixels of an MxN image is turned into PxQ image blocks where the size of the block is determined by input.

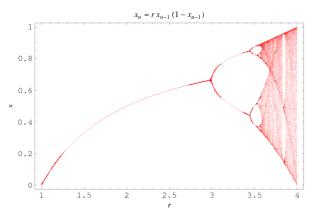


Figure 1

As we can see in figure 1 after a certain value the function starts giving very chaotic values which are hard to predict. It is also highly sensitive to the intial values r and t, even a minor change might give a completely different pseudo random sequence. Compared to other pseudo random functions such as lorentz function this is simpler and easier to compute.

The benefits of chaotic encryption technology include simple implementation, robustness, fast encryption, and high security. However, although chaotic encryption technology has many advantages, it also has a number of deficiencies. For example, at present, most chaotic encryption algorithms confuse the single image pixel value or location, but the utilization of only one of the two strategies (we use only confusion) do not ensure high security for the image, and thus it is easy for attackers to crack an encrypted image by simply using the pixel comparison method. In 1994, Adleman first introduced DNA computing into the encryption field, which created a new stage of information processing. DNA encryption is a new frontier and is presently at the forefront of international cryptography research. DNA molecules harness massive parallelism and have low energy consumption and high storage density. Therefore, image encryption algorithms based on DNA computing

possess unique advantages that the traditional cryptographic algorithms do not have. However, using only DNA encoding to encrypt images is not secure. Therefore, we combine chaos encryption technology and image encryption based on DNA computing to solve the hidden insecurity problems existing when images are confused using the chaotic encryption technology. First, we confuse the digital image pixels using the chaotic encryption technology. We then diffuse the confused pixels using DNA encoding to obtain the final image.

In summary, our study successfully combines chaotic encryption technology and DNA coding techniques in a method that has been verified via a large number of experiments and security analyses to prove the security and rationality of the algorithm.

DNA Encoding and Chaotic Maps

DNA Encoding and Complementary

Rule. DNA sequencing is the process used to map the nucleotide sequence forming a strand of DNA. Four bases, adenine (A), thymine (T), guanine (G), and cytosine (C) form the building blocks of genetic code. "A" bind with "T" and "G" binds with "C" [17]. We know that every digital image pixel can be expressed by 8- bit binary numbers [18]. Because the binary numbers "0" and "1" are complementary, "00" and "11" and "01" and "10" are also complementary. If we use the four deoxynucleotides "A," "T," "G," and "C" to represent the binary numbers "00," "11," "01," and "10," respectively, then each pixel can be encoded into a string of nucleotides. For example, the grey value of a digital image pixel is 228, and the binary corresponding to this value is "11100100." According to the above rules, the string of

nucleotides that corresponds to this binary is "TCGA." There are 24 types of combinations for the four nucleotides. However, only eight coding combinations are suitable for the principle of complementarity. These rules are summarized in Table 1.

TABLE 1: The rules of DNA encoding.

1	2	3	4	5	6	7	8
00-A	00-A	00-C	00-C	00-G	00-G	00-T	00-T
01-C	01-G	01-A	01-T	01-A	01-T	01-C	01-G
10-G	10-C	10-T	10-A	10-T	10-A	10-G	10-C
11-T	11-T	11-G	11-G	11-C	11-C	11-A	11-A

4. ALGORITHM

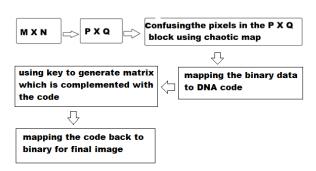


Figure 2: Block diagram of proposed model

Suppose we take a MxN image, we divide it into its RGB values hence obtaining three binary matrices (MxN). The steps are as follows

INPUT: Image I, the initial values of logistic fucntion t and x. The value p in range 0-30 for confusion and the key for image diffusing using DNA mapping.

OUTPUT: the scrambled image

Step 1. Convert the original image *I* into 3 two-dimensional matrices called R,G and B. Let R and C be used to record the rows and columns of the 2-D matrix.

Step 2. Input the range of confusion p(0-30) which is used to create a chaotic index using logistic mapping function, first for rows and then columns which is used to change the pixels row-wise and columnwise respectively (for all three matrices). Thus, we can obtain a confused image *I*'.

$$X_{n+1} = \mu X_n (1 - X_n).$$

Logistic Mapping Expression

Step 3. Convert image I' into a binary two-dimensional matrix I''. The size of I'' is $M \times N$ rows and eight columns. We use the selected DNA encoding rule from the table (Table 1) to convert I'' into a DNA encoding matrix with $M \times N$ rows and four columns. Finally, convert this matrix into a one-dimensional DNA coding sequence X, whose size is $M \times N \times 4$.

Step 4. The key is used to generate a KEY matrix of the same format and size as the DNA matrix. This KEY is then used to complement the DNA matrix. This is much like the XOR function except with 4 characters which is later mapped using the rule in figure 3.

```
# DNA xor

dna["AA"] = dna["TT"] = dna["GG"] = dna["CC"] = "A"

dna["AG"] = dna["GA"] = dna["TC"] = dna["CT"] = "G"

dna["AC"] = dna["CA"] = dna["GT"] = dna["TG"] = "C"

dna["AT"] = dna["TA"] = dna["CG"] = dna["GC"] = "T"
```

Figure 3

Ex.
$$dna[A] + key[A] = "A"$$

Step 5. Convert the DNA encoded matrix back to a binary matrix of size MxN for all three matrices. Output the final image.

5. Modules Involved:

In terms of functions defined in python code:

- decompose_matrix decomposing the image into RGB matrices
- dna_encode generating DNA matrix
- dna_decode generating binary matrix
- recover_image saves final image from matrices
- key_matrix_encode generate encoded key matrix from key
- xor_operation xor operation on DNA matrix using key matrix
- factor generated factors of a numerical value
- keygen generates the index for confusion using logistic function
- shuffle confuses the pixels using index

SCRIPTS

login – GUI for interface

encryption – pixel confusing/scrambling using value and logistic mapping function

dna – DNA encryption

libraries used in python - tkinter, cv2, numpy, PIL

Contributions:

- 1. Mugesh login (script), encryption(script)
- 2. Nikhil login(script), image preprocessing
- 3. Archit DNA encoding

CODE

login(script)

```
def login(self):
def encrypt(self):
```

```
root=Tk()
m=login(root)
root.mainloop()
```

encryption(script)

```
import cv2
import numpy as np
           key1.append(i)
def keygen(x,r,size):
       key.append(x)
       index.append(i)
                key[i], key[j]=key[j], key[i]
   m, n = b.shape
   print(B1.shape, b.shape, x)
   key = keygen(0.01, 3.95, x)
   key = keygen(0.01, 3.95, y)
```

```
path = filedialog.askopenfilename()
    img = np.zeros((values.shape[0], values.shape[1]), dtype=np.uint8)
img = Image.open(path)
```

dna(script)

```
from PIL import Image
import cv2
import numpy as np
dna = {}
dna["00"] = "A"
dna["01"] = "T"
dna["10"] = "G"
dna["11"] = "C"
dna["11"] = [0, 0]
dna["T"] = [0, 1]
dna["G"] = [1, 0]
```

```
def xor operation(b, g, r, mk):
   m, n = b.shape
   m, n = b.shape
```

```
Mk enc = Mk enc.astype(str)
def decompose_matrix(iname):
def recover_image(b, g, r, iname):
```

6. OUTPUT

If we run the login script a GUI opens

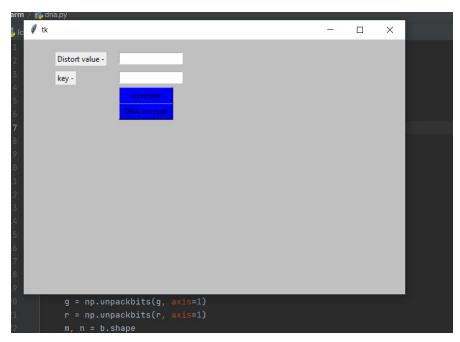


Figure 4.1

First we enter the distort value (0-30) and click scramble where a dialog box opens to select the image.

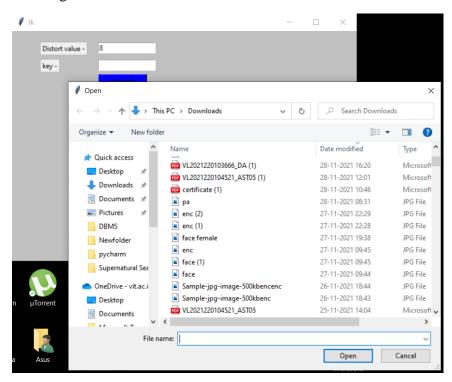


Figure 4.2

We then select an image which then returns the scrambled output.

Before scrambling-

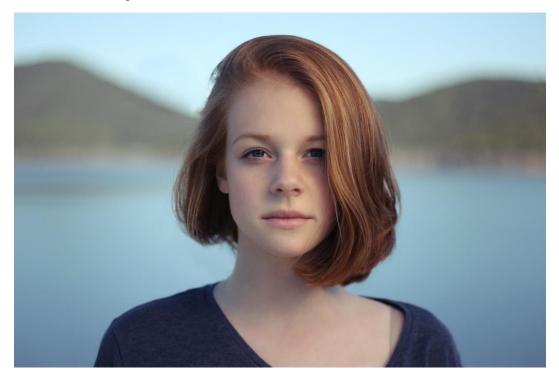


Figure 5.1

After scrambling –

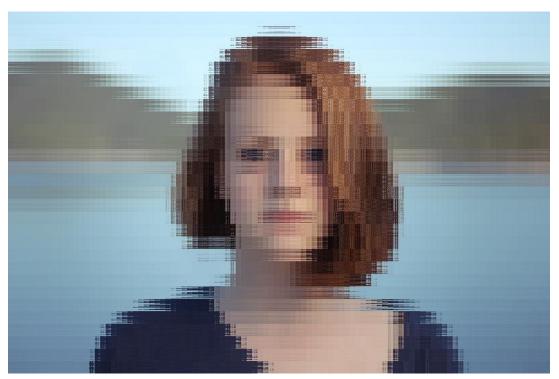


Figure 5.2

Now you enter a key for DNA encrypting which encrypts the scrambled image

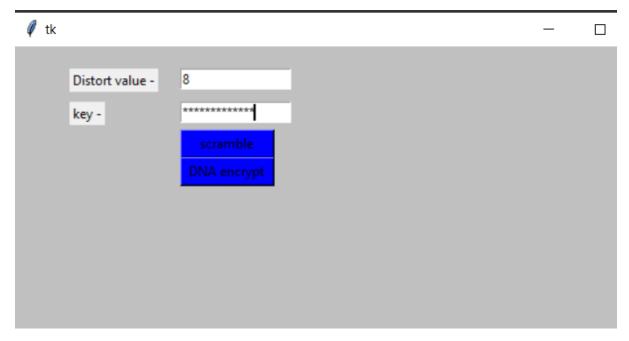


Figure 4.3

After DNA encrypt -

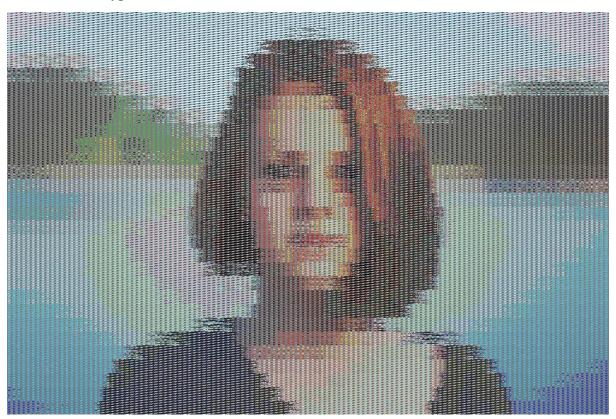


Figure 5.3

Test2



Scramble = 5



DNA encrypt = Joethiswar11@



Figure 6.1

Figure 6.2

Figure 6.3

Test 3



Scramble=5



DNA encrypt= Joethiswar11@

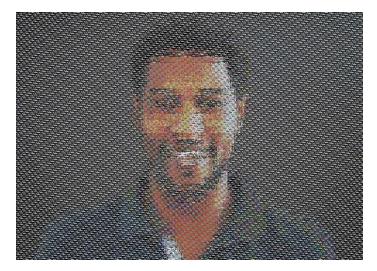


Figure 7.1

Figure 7.2

Figure 7.3

COMPARISION

histogram depicting the tonal distribution of the images before and after the algorithm CODE



Figure 6.1 Figure 6.2

Figure 6.1 and 6.2 show the amount of pixels in the intensity from 0-255. We can see than DNA encoding has drastically altered the pixel values.

We use an online image comparision tool https://www.betafaceapi.com/demo.html to compare the images for identifying utility.

Search RESULT



Original Face



Figure 7

It shows that the image after processing is 68.5% identical to the original one.

It also provides 21 facial points which can be used to describe the face.

ORIGINAL

5oclock shadow: no, age: 22 (60%), arched eyebrows: yes (3%), attractive: yes (82%), bags under eyes: no (96%), bald: no (79%), bangs: no, beard: no, big lips: no (45%), big nose: no, black hair: no, blond hair: no (10%), blurry: no (67%), brown hair: yes (65%), bushy eyebrows: no (82%), chubby: no, double chin: no, expression: neutral (79%), gender: female (98%), glasses: no, goatee: no, gray hair: no (94%), heavy makeup: yes (31%), high cheekbones: no (64%), mouth open: no, mustache: no, narrow eyes: no (59%), oval face: yes (74%), pale skin: yes (36%), pitch: -10.02, pointy nose: no (25%), race: white, receding hairline: no (27%), rosy cheeks: no, sideburns: no, straight hair: no (7%), wavy hair: yes (12%), wearing earrings: no (49%), wearing hat: no, wearing lipstick: yes (58%), wearing necklace: no (51%), wearing necktie: no (85%), yaw: -5.05, young: yes (94%),

PROCESSED

5oclock shadow: no (90%), age: 39 (60%), arched eyebrows: no, attractive: no (75%), bags under eyes: no (15%), bald: no (61%), bangs: yes (48%), beard: no (90%), big lips: no (53%), big nose: no (24%), black hair: no, blond hair: yes (14%), blurry: yes (19%), brown hair: no (13%), bushy eyebrows: no, chubby: no (53%), double chin: no (66%), expression: neutral (74%), gender: female (5%), glasses: no (97%), goatee: no, gray hair: yes (8%), heavy makeup: no (83%), high cheekbones: no (73%), mouth open: no (68%), mustache: no (92%), narrow eyes: yes (0%), oval face: no (54%), pale skin: yes (77%), pitch: -6.8, pointy nose: no (65%), race: white, receding hairline: no (56%), rosy cheeks: no, sideburns: no, straight hair: yes (27%), wavy hair: no (55%), wearing earrings: no (68%), wearing hat: no (54%), wearing lipstick: no (68%), wearing necklace: no (54%), wearing necktie: no (46%), yaw: 3.76, young: yes (27%),

Changed characteristics

age, arched eyebrow, attractive, bags under eyes, bangs, blonde hair, blurry, grey hair, heavy makeup, narrow eyes, oval face, straight hair, wavy hair, lipstick

Same characteristics

Bags under eyes, bald, beard, big lips, big nose, black hair(no), bushy eyebrows(no), chubby, double chin(no), expression(neutral), gender, no glasses, no goatee, pointy nose(no), no sideburns, pale skin(yes), race, receding hairline(no), not wearing hat, young(yes), not wearing neckloace or necktie.

7.DISCUSSION

Researchers at Japan's National Institute of Informatics (NII) have found that fingerprints can be easily recreated from photos taken up to three metres away without the need for advanced technology. Many phones nowadays use facial recognition for unlocking. It's obvious our images hold many potential data which could be used to target us.

Key Space and Sensitivity Analysis: Any chaotic system is sensitive to the initial values. To make the encryption algorithm highly secure, the key space should be large enough to make any brute-force attack ineffective. Here, all the keys are from the process of confusing and diffusing the pixels. Our encryption algorithm actually does have some of the following secret keys:

- (1) the initial values of the logistic map, a0 and b0, and the parameters μa and μb ;
- (2) DNA coding and complementary rules and key.

Gray Histogram Analysis. Following statistical analysis of the original and anonymized images, we constructed grayscale histogram analysis of figure 5.1 and 5.3. Comparing the histograms the pixel values of the original image and anonymized image are highly different which makes statistical attacks difficult.

Hence its imperative that while releasing the image for mining it should be in a format that individuals privacy mustn't be compromised.

8.CONCLUSION

In this paper, we proposed a novel confusion/diffusion algorithm for image encryption. First, we exchanged the pixel positions of rows and columns of the digital image according to a chaotic index based on the logistic chaotic map to confuse the image pixels. Then, we encoded each of the pixels that had been confused into four nucleotides. A key was used to further encrypt the image by and xor function on the DNA code. Finally, we converted the two dimensional matrix obtained into a distorted image. This algorithm provides a safe was for images to be transformed in a way that they still retain some utility while also protecting privacy. The combination of the algorithms also has been proven to withstand brute force attacks as multiple combinations of values and two methods are used to obtain the final processed image. Therefore, the scheme is reliable enough to be applied in image processing for utility.

9. REFERENCES

1. Y. Liu, X. Tong, and S. Hu, "A family of new complex number chaotic maps based image encryption algorithm," *Signal Processing: Image Communication*, vol. 28, no. 10, pp. 1548–1559, 2013

- 2. A. Kadir, A. Hamdulla, and W.-Q. Guo, "Color image encryption using skew tent map and hyper chaotic system of 6th-order CNN," *Optik*, vol. 125, no. 5, pp. 1671–1675, 2014
- 3. J. W. Yoon and H. Kim, "An image encryption scheme with a pseudorandom permutation based on chaotic maps," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 12, pp. 3998–4006, 2010.
- 4. Q. Zhang, L. Guo, and X. Wei, "Image encryption using DNA addition combining with chaotic maps," *Mathematical and Computer Modelling*, vol. 52, no. 11-12, pp. 2028–2035, 2010
- 5. Q. Zhang, Q. Wang, and X. Wei, "A novel image encryption scheme based on DNA coding and multichaotic maps," *Advanced Science*

- *Letters*, vol. 3, no. 4, pp. 447–451, 2010.
- 6. S. H. Jiao and R. Goutte, "Code for encryption hiding data into genomic DNA of living organisms," in *Proceedings of the 9th International Conference on Signal Processing (ICSP '08)*, pp. 2166–2169, Beijing, China, October 2008
- 7. L. Liu, Q. Zhang, and X. Wei, "A RGB image encryption algorithm based on DNA encoding and chaos map," *Computers and Electrical Engineering*, vol. 38, no. 5, pp. 1240–1248, 2012.
- 8. M. François, T. Grosges, D. Barchiesi, and R. Erra, "A new image encryption scheme based on a chaotic function," *Signal Processing: Image Communication*, vol. 27, no. 3, pp. 249–259, 2012.