Research on intelligent recognition and encryption algorithm of network big data image

Jubao Qu*
School of mathematics and computer, Wuyi University, Fujian 354300, China
The Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian Education Institutions (Wuyi University), Wuyi University Wuyishan, Fujian 354300, China

*Corresponding author: jubaoqu@wuyiu.edu.cn

Abstract. In view of the identification and encryption of massive network information in the era of big data, this paper proposes to use intelligent image recognition technology to analyze the required information, then classify it according to the big data clustering method, and use pseudo pixel encryption algorithm for multiple encryption. The simulation results show that the algorithm has high recognition rate, good encryption effect, strong anti-interference ability and good ability to protect network image information.

Keywords: Big data, intelligent recognition, encryption, image.

1. Introduction
With the advent of 5G era, network transmission is speeding up day by day, the world's demand for information is also growing, especially for network image information. How to ensure that a large number of network images can be correctly identified and encrypted transmission has become the main problem to be solved in the field of Internet. However, most image encryption methods cannot modify the transformed pixels when encrypting Internet image resources, which makes the implementation of encryption of Internet users' image resources fall into a bottleneck, Reference [3] uses two-dimensional random matrix scrambling transformation method to encrypt network image. Their methods can be well implemented in a certain specific environment, but there are some limitations, the degree of intelligence is not high, and the encryption process is time-consuming and cumbersome.

As an important research field of artificial intelligence, computer intelligent image recognition technology provides a reliable method for effective recognition of various kinds of images, The image that needs to be recognized can be input into the computer system, and then the computer will match the image itself, extract the characteristics of the image, and make specific classification [4]. In the process of intelligent image recognition, the preprocessing link is image input, which is a key link, directly related to the recognition effect of the image. The collected image is input into the computer. At this time, the recognition system will separate the background from the image according to the characteristics of the image itself, and refine the image binarization as a whole, so as to ensure the processing speed and efficiency of the image in the later stage, so as to ensure the authenticity of the image to the maximum extent, reduce the false components[5], and complete the image recognition with the help of relevant
values Therefore, this paper proposes a network big data image intelligent recognition and encryption algorithm, through the network image intelligent recognition, and according to the big data clustering method classification[6], uses the pseudo pixel encryption algorithm to encrypt. Experimental results show that the proposed method has high recognition rate, good encryption effect, strong anti-interference ability and good ability to protect network image information.

2. Algorithm of computer intelligent image recognition

2.1. Improved Hu moment invariants algorithm

In the massive information on the Internet, recognition and encryption of important images is a new technology in image processing. How to use artificial intelligence means and computer intelligent image recognition technology to identify different targets and objects has become the key technology. At present, there are two commonly used algorithms for intelligent image recognition: Hu moment invariants, D-S evidence reasoning, etc [7]. By improving them, we can obtain very good image recognition effect. Hu invariant moment is one of the earliest algorithms for intelligent image recognition. It is composed of seven variable values by the nonlinear combination of low order normalized central moments.

When the two-dimensional digital image is represented by \( \mathcal{R}(x, y) \), the order moment of the image can be defined as follows [8]:

\[
\Psi_{ij} = \sum_{x} \sum_{y} x^i y^j \cdot \mathcal{R}(x, y), \quad (i, j = 0, 1, 2, m)
\]

At the same time, the central moment of order \((i+j)\) can be defined as follows:

\[
\eta_{ij} = \sum_{x} \sum_{y} (x - x_0)^i (y - y_0)^j \cdot \mathcal{R}(x, y)
\]

Where, \( x_0 = \Psi_{10}/\Psi_{00} \), \( y_0 = \Psi_{01}/\Psi_{00} \), for two-dimensional digital images, its centroid coordinates are \( \mathcal{R}_0(x_0, y_0) \), in which \( x_0 \) represents the centroid of image gray level in the horizontal direction, and \( y_0 \) represents the centroid of image gray level in the vertical direction. The normalized central moments of order \((i+j)\) are defined as follows:

\[
\varphi_{ij} = \eta_{ij} / \eta_{00}
\]

In the formula, \( t = \frac{(i+j+2)}{2}, i+j = 2, 3, m \).

2.2. Improved D-S evidence reasoning algorithm

Computer intelligent image recognition belongs to the category of multi-source information processing. In this kind of information processing, data fusion is a new technology, and D-S evidence reasoning is one of the main research methods. It is based on the non-empty set \( \Phi \), also known as the discrimination framework. Through this framework, we can describe the set of elements that constitute the hypothesis space. The requirement of the framework is that the elements should be mutually exclusive. The power set composed of \( \Phi \) subsets can be represented by \( 2^\Phi \). Then the trust assignment function is defined as follows in the power set:

\[
\Gamma(S) \rightarrow [0, 1]
\]
Where \( s \) denotes any subset of the framework, and \( \Gamma(S) \) represents the support degree of the evidence for the occurrence of proposition \( S \). Moreover, it needs to satisfy the following conditions.

\[
\Gamma(\Phi) = 0 , \quad \sum_{S \in s} \Gamma(S) = 1
\]  

(5)

3. Image big data classification encryption algorithm

Aiming at the problem of large amount of network image information, in order to improve the speed of computer intelligent image recognition, the idea of big data can be used to cluster a large number of image data points according to the big data classification method. First of all, we need to obtain the clustering center of big data, and classify the image big data as follows.

\[
Z(\mu) = \sum_{i,j=1}^{n} \mu_{ij} \cdot \rho(X_i, Y_j)
\]  

(6)

Among them, \( \rho(X_i, Y_j) \) is the distance between a pair of big data \( (X_i, Y_j) \), and \( \mu_{ij} \) is the clustering threshold. When it is greater than \( \mu_{ij} \), clustering will stop.

After classifying the image data, the data privacy encryption is carried out. The secret key of data encryption is obtained by using the following formula to encrypt the data.

\[
\Omega(p, q) = (p - 1) \cdot (q - 1) / Z(\mu)
\]  

(7)

Where \( p \) is the public key of data encryption and \( Q \) is the private key of data encryption.

4. Improved pseudo pixel encryption algorithm

In the process of encrypting the Internet user's image resources, the target pixel is selected according to the key before the image resource is encrypted. The bit data of the target pixel is embedded into other pixels, and the image is encrypted. In the resource encryption domain, the information to be embedded is formed into pseudo pixels, and the target pixel is replaced by the pseudo pixel.

Let \( \mathcal{R}(m, n) \) represent the original image with the size of \( m \times n \), and divide it into multiple image blocks with the size of \( i \times i \). \( M \) and \( N \) are integral times of \( I \). The image block is divided into multiple T-type pixel groups with every four pixels as a group, as shown in Fig. 1. Take two large prime numbers \( P \) and \( Q \), \( m = p \times q \). The following formula is used to select the target pixel according to the key before image encryption.

\[
H_{ij}(t) = \frac{B'_{ij}(t) + H_{ij}(t)}{k_{t}} \oplus \frac{m \times n}{k_{t}} \otimes T \otimes \{p, q\} \otimes \mathcal{R} \otimes \{i \times i\}
\]  

(8)

Where \( B'_{ij}(t) \) represents the edge pixel, \( k_{t} \) represents the center pixel, \( k_{t} \) represents the encryption key, and \( H_{ij}(t) \) represents the number of pixel groups.
In the process of encrypting the image information of Internet users, the pseudo pixel encryption algorithm is used to give the transformation key, and the adjacent pixels of the image to be encrypted are transformed into integers. The transformed pixel groups are found in the encrypted image by using the encryption key, and the pixels used for information embedding in each pixel group are selected by the embedding key.

5. Experiment and simulation
In this paper, we take Lena, baboon and pepper, which are 256 × 256 gray-scale images, as examples, and use matlab 2014a software to carry out simulation experiments. Experimental environment: PC configuration: Intel (R) core (TM) i5-4590 CPU @ 3.30 GHz 3.30 GHz, memory 4GB, 64 bit windows7 operating system. In the experiment, three gray images, Lena, baboon and barbarn, are selected, and the pixel size is 512×512.

5.1. Histogram analysis of encrypted image

![Figure 1. Schematic diagram of image T-type grouping](image)

![Figure 2. Histogram comparison before and after encryption](image)
In order to analyze the change of image pixels before and after encryption, histogram analysis is needed. Figure 2 shows the images baboon, Lena and pepper. The histogram before and after encryption can be seen from the figure that the histogram of plaintext image fluctuates obviously with the change of pixel gray value, which directly reflects the distribution characteristics of image pixel value; the histogram display of cipher text image is very smooth, indicating that the gray value of cipher text image is evenly distributed in the value range of [0-255], which can well hide the statistical information of cipher text image; strong resistance to statistical analysis.

5.2. Comparison of image encryption and decryption results
Three images, baboon, Lena and pepper, are tested by the algorithm in this paper. Figure 3 is the comparison of plaintext image, ciphertext image and decryption image. It can be seen from Figure 3 that the encrypted image has become disordered and completely different from the plaintext image in terms of vision. It shows that the encryption visual effect of this algorithm is good, and the decrypted image is completely correct, which shows that the algorithm designed in this paper can correctly realize image encryption and decryption. The decrypted image is consistent with the plaintext image visually, which shows the feasibility of this method.

5.3. Analysis of SNR and embedding rate of encrypted image
Finally, the SNR and embedding rate of the proposed method are calculated. The statistical results are shown in Table 1. From the experimental results in Table 1, it can be concluded that the proposed method for Internet user image resource encryption has good signal-to-noise ratio and embedding rate, which
fully guarantees the comprehensive effectiveness of the proposed method for resource encryption security protection.

### Table 1. Analysis of RMS error and PSNR of image

| Image name | SNR(dB) | Information embedding rate (%) |
|------------|---------|-------------------------------|
| Baboon     | 53.65   | 98.7                          |
| Lena       | 62.58   | 96.2                          |
| Peppers    | 64.24   | 97.9                          |

5.4. **Correlation analysis of neighboring pixels in cipher text image**

In order to test the three-dimensional correlation of adjacent pixels in the encrypted image, 900 pixels and its adjacent pixels in the horizontal, vertical and diagonal directions were randomly selected from the plaintext image and cipher text image of the three images respectively, and the correlation coefficients of each image in the horizontal, vertical and diagonal directions were calculated. Table 2 shows that the correlation between adjacent pixels of the original image in the horizontal, vertical and diagonal directions is very strong, and the correlation coefficient of the adjacent pixels in the three directions of the cipher text image is very small, and the correlation coefficient has tended to 0, which indicates that the statistical characteristics of the plaintext image have been well spread into the cipher text image, and have good privacy characteristics.

### Table 2. Comparison table of correlation coefficient between plaintext image and encrypted image

| Image name | Horizontal correlation coefficient | Vertical correlation coefficient | Diagonal correlation coefficient |
|------------|----------------------------------|---------------------------------|---------------------------------|
|            | Plain-text image | Encry-pTed image | Plain-text image | Encry-pTed image | Plain-text image | Encry-pTed image |
| Baboon     | 0.9365 | 0.0075 | 0.9687 | -0.002 | 0.8641 | 0.053 |
| Lena       | 0.7659 | -0.005 | 0.9635 | 0.027 | 0.9135 | -0.027 |
| Peppers    | 0.9761 | 0.0098 | 0.9236 | 0.016 | 0.9642 | 0.017 |

5.5. **Information entropy analysis**

Information entropy reflects the uncertainty of an information. Image information entropy reflects the distribution measurement of image gray value. The more uniform the gray value distribution is, the greater the image information entropy is, and otherwise, the smaller the information entropy is. The calculation formula is as follows:

$$K(u) = -\sum_{i=1}^{n} I(x_i) \cdot \log_2 I(x_i)$$  \hspace{1cm} (9)

According to formula (9), the information entropy of encrypted test image is calculated. As shown in Table 3, the information entropy of encrypted image is above 7.999, which is very close to the theoretical value of 8. It shows that the gray value distribution of the encrypted image is very uniform, which shows that the algorithm has a good ability to resist statistical analysis attacks.

### Table 3. Information entropy analysis table of encrypted image

| Image name     | Baboon | Lena |Peppers |
|----------------|--------|------|--------|
| Plaint text image | 7.2698 | 7.5246 | 7.7023 |
| Cipher text image | 7.9997 | 7.9993 | 7.9991 |
5.6. Analysis of fixed point ratio of encrypted image

Through the comparative analysis of image fixed point, the smaller the fixed point ratio of encrypted image, the greater the difference between cipher text image and plaintext image, the better encryption effect. However, the fixed point ratio only reflects the change of pixels, but fails to reflect the change of pixel gray value. If the gray value of all pixels in the image is increased or decreased by a constant, the visual information of the image is basically changed. Therefore, the fixed point ratio is often combined with the average change value of gray level to evaluate the image encryption effect. From this index, more than 99% of the gray values of the image points have changed, which has a good encryption effect.

6. Conclusion

Using the improved Hu invariant moment algorithm, combined with the image big data classification encryption algorithm and the improved pseudo pixel encryption algorithm, the massive network information in the era of big data is identified and encrypted. The intelligent image recognition technology is used to analyze the required information, and then classified according to the big data clustering method. The pseudo pixel encryption algorithm is used for multiple encryption. The simulation results show that the algorithm has high recognition rate, good encryption effect, strong anti-interference ability, and has a good ability to protect network image information, and has certain adaptive value.

Acknowledgments

This work has been supported by the following funds: The Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian Education Institutions(Wuyi University), Fujian Provincial offline and social practice first-class undergraduate course construction project (sjke2020006), Fujian University student innovation and entrepreneurship training program project (201910397047), provincial first-class undergraduate professional construction in 2019 (sjzy2019003).

References

[1] Zhou N R, Zhang A D, Zheng F. Novel image compression–encryption hybrid algorithm based on key-controlled measurement matrix in compressive sensing. Optics and Laser Technology, 2014, pp. 152 - 160.
[2] Gong L H, Deng C Z, Pan S M. Image compression-encryption algorithms by combining hyper-chaotic system with discrete fractional random transform. Optics and Laser Technology, 2018, pp. 48 - 58.
[3] Do T T, Gan L, Nguyen N H, et al. Fast and efficient compressive sensing using structurally random matrices. IEEE Transactions on Signal Processing, 2012, pp. 120 - 126.
[4] S. Yu, W. K. S. Tang. Generation of X-scroll Attractors in a Two-Port RCL Network with Hysteresis Circuits. Chaos Solitons & Fractals, 2009, pp. 821 - 830.
[5] G. Zhang, F. Wang. A Novel Multi-scroll Chaotic Generator: Analysis, Simulation, and Implementation. Chinese Physics B, 2018, pp. 587 - 591.
[6] Y. Liu, X. Tong, J. Ma. Image Encryption Algorithm Based on Hyper-Chaotic System and Dynamic S-box. Multimedia Tools & Applications, 2016, pp. 706 - 710.
[7] A. R. Elsonbaty, A. M. A El-Sayed. Further Nonlinear Dynamical Analysis of Simple Jerk System with Multiple Attractors. Nonlinear Dynamics, 2017, pp. 1106 - 1109.
[8] G. Maddodi, A. Awad, D. Awad, M. Awad, B. Lee. A new Image Encryption Algorithm Based on Heterogeneous Chaotic Neural Network Generator and DNA Encoding. Multimedia Tools and Applications, 2018, pp. 561 - 565.