Publication Date: 20 August 2024
*Archs Sci. (2024) Volume 74, Issue S2 Pages 199-208, Paper ID 2024s223.*
*https://doi.org/10.62227/as/74s223*

# Research on Image Encryption Algorithm Based on Generative Neural Network

**Guo Li[1,*], Yi Qin[1] and Minghua Wang[2]**
[1]College of Intelligent Manufacturing and electrical engineering, Nanyang Normal University, Nanyang, Henan, 473000, China.
[2]Shandong Gete Aviation Technology Co., Ltd, Jinan, Shandong, 250000, China.
Corresponding authors: (e-mail: airforce1980@126.com).

**Abstract** The security of digital image transmission is related to the national economy and people's livelihood, and the traditional text encryption algorithm performs poorly on images. Based on this, this paper designs the image encryption model, optimizes the generator and discriminator network structure based on the residual network model, and improves the generative adversarial network. Logistic chaotic system is utilized to construct a new GAN key generation model, and it is applied to the field of image encryption. On this basis, the Knuth-Durstenfeld algorithm is improved to complete the encryption of images from three processes: pixel value replacement, image disruption, and image diffusion, and applied research is carried out. The results show that the histograms of Lena and Peppers plaintext images show irregular statistical characteristics, and the gray value distribution pattern of the histograms of ciphertext images has no big ups and downs and tends to be stable and smooth. The correlation coefficients of adjacent pixels of ciphertext image are close to the ideal value of 0, and the image information is more difficult to be analyzed and deciphered. Under the noise attack, the PSNR values of the ciphertext image do not differ much, which can prove that the ability of anti-noise attack is good. The encryption algorithm based on generative adversarial network shows excellent performance on image encryption and protects the information security of digital images.

**Index Terms** image encryption, generative adversarial network, key generation model, Knuth-Durstenfeld algorithm.

## I. Introduction

With the rapid advancements in Internet and multimedia technologies, multimedia communication has emerged as a pivotal means for individuals to exchange information among themselves [1]. Concurrently, as people engage in the exchange of diverse information via the Internet, new requirements for information security in safeguarding confidential information during network transmission have arisen, gradually attracting research focus and attention [2-3].

Serving as a pivotal medium for information transmission, digital images occupy a significant role in facilitating the exchange of information. In the realm of network transmission and storage, numerous private and confidential images, encompassing patient medical records, as well as military and national defense-related sensitive imagery, are prevalent [4-5]. Should these images fall into unauthorized hands through illicit means, they pose a grave risk of exposing the user's personal vital information, and in extreme cases, could potentially jeopardize national security. Consequently, the security of image information has garnered widespread attention among scholars [6-7].

Traditional image encryption algorithms commonly perceive digital images as sequences of binary data, subsequently encrypting these sequences using established data encryption methodologies, notably the Data Encryption Standard and the Advanced Encryption Standard [8]. Nevertheless, given the extensive information content inherent in digital images and the substantial redundancy among adjacent pixels, traditional encryption approaches necessitate substantial computational resources, with their security foundation resting on the intricacy of the underlying mathematical computations. In contrast, deep learning, a sophisticated machine learning paradigm, has garnered widespread application across diverse domains, encompassing target detection, image synthesis, speech recognition, and more. Notably, deep learning excels in its proficiency for intricate feature extraction and learning capabilities, enabling it to discern intricate patterns and characteristics within images [9-10].

In recent years, the realm of image encryption has witnessed the emergence of deep learning technology as a promising approach. The primary limitation of traditional image encryption algorithms stems from their reliance on a predefined set of equations, which contrasts with deep learning-based encryption methods. These latter algorithms, owing to their extensive computational models and intricate processes, exhibit an enhanced capability to withstand at-

tacks, offering heightened encryption strength and improved security. Furthermore, they facilitate increased efficiency in image encryption, thereby accommodating a diverse array of scenarios and application demands [11].

In recent years, the rapid development of computer technology has led to a heightened concern among an increasing number of individuals regarding the issue of information security. Concurrently, the extensive utilization of digital images has necessitated the transmission of vast amounts of image information over networks, posing a risk of theft by unauthorized users during storage or transmission, ultimately resulting in information leakage. To address this challenge, image encryption emerges as a pivotal means for safeguarding the security of digital image information [12]. Jia, S. et al. have devised an encrypted image retrieval algorithm centered on scene understanding, which not only achieves a commendable level of retrieval accuracy but also enhances the security of users' image privacy [13]. Additionally, Nguyen, T. S. et al. have proposed a cryptographic image reversal algorithm that utilizes the Absolute Moment Block Truncation Coding (AMBTC) algorithm as the fundamental architecture for reverse data hiding in encrypted images. In performance simulation tests, their algorithm has consistently outperformed traditional methods [14]. Tang, Z. et al. presented the significance of reversible data hiding (RDH) in safeguarding data information security and evaluated the efficacy of their proposed RDH algorithm when applied to encrypted images in a standardized benchmark image data test library, demonstrating exceptional embedding capacity [15]. Meanwhile, Ye, G. et al. formulated a novel multiple illusions image encryption method, incorporating compression-aware theory and Schur decomposition concepts, tailored to meet the stringent requirements of image privacy security in communication environments [16]. Furthermore, Li, X. et al. devised a hybrid mapping image encryption algorithm, integrating random arithmetic strategies from various randomized DNA encoding techniques, and conducted thorough experiments which revealed the algorithm's capacity to withstand typical information attacks [17].

In the realm of multimedia information, image information has emerged as a pivotal means of human expression due to its vivid depiction. However, it encompasses classified image data, such as military equipment deployment diagrams and confidential drawings pertaining to industries like business and construction, which necessitate encryption prior to transmission over networks. Hu, G. F. refined the concept of quantum image flexible representation and devised an advanced quantum image encryption algorithm, thereby bolstering the privacy and security of image transmission [18]. Abd-Alhameed, R. A. introduced a capacitive hyperchaotic system grounded in four-dimensional fractional order memory theory, incorporating an encryption system for image transmission within the framework to ensure efficient and stable realization of image encryption protection [19]. Farwa, S. et al. examined the viability of a scheme that integrates the algebraic substitution box effect and the Arnold transform dislocation effect for image encryption. Through experimental and statistical analysis, they demonstrated the scheme's ability to achieve image encryption and safeguard against cyber-attacks [20]. Kumar, P. et al. pioneered an algorithm for the fusion generation of medical and visual images, leveraging the principle of non-downsampled shearlet transform for structural composition. The images generated by this algorithm exhibited superior definition, precision, and other desirable qualities in functional testing [21].

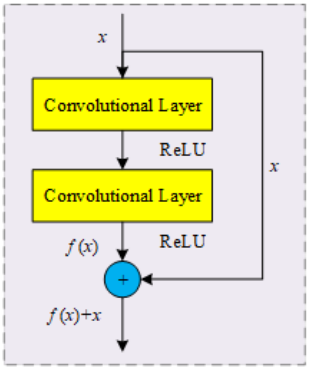


Figure 1: Residual module

In this paper, we first propose a generative adversarial network (GAN) framework grounded on the residual network model. Subsequently, we optimize both the generator network and the discriminator network structures individually to enhance the stability of the training process. Furthermore, we incorporate the Wasserstein distance to address the limitations of the loss function. Subsequently, we construct a novel GAN-based key generation model tailored to the requirements of image encryption security, leveraging the Logistic chaotic system.To this end, we develop the overall architecture of an image encryption algorithm by refining the Knuth-Durstenfeld algorithm. This encryption approach encompasses three primary facets: pixel value substitution, image disruption, and image diffusion. Ultimately, we conduct experimental evaluations of the devised encryption system, subjecting the ciphertext images to a rigorous analysis encompassing histogram analysis, neighboring pixel correlation assessment, robustness testing, and resistance analysis against differential attacks. These analyses aim to comprehensively evaluate and demonstrate the encryption performance of the proposed algorithm on image data.

## II. Image Encryption Model Construction

### A. Residual network modeling

The basic structure of ResNet is similar to the convolutional neural network, the difference is that ResNet adds a special residual module, and the basic structure of the residual module is shown in Figure 1.

Residual module Figure 1 shows that $x$ skips two convolutional layers from the right, this structure is called "jump connection", the residual module is constructed with this connection as the core. Assuming that the mapping of the

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

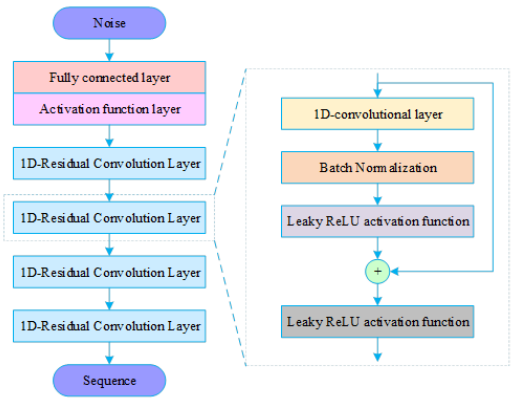


Figure 2: The gencrator structure of 1D-RGAN

two convolutional layers in Figure 1 is $f(x)$, if the "jump connection" on the right is not added, the input $x$ will be mapped by $f(x)$ to get $H(x)$, and the output of the network is shown in (1).

$$H(x) = f(x). \tag{1}$$

After adding "jump connection", the output is shown in (2).

$$H(x) = f(x) + x. \tag{2}$$

In the deep network, it is easy to appear the phenomenon of decreasing effect, at this time, the input can be made equal to the output, skipping the mapping process of decreasing effect, i.e., $H(x) = x$. For the ordinary convolutional network, whose output is $H(x) = f(x)$, in order to make $H(x) = x$, it is necessary to train the neural network to make $f(x) = x$, which is difficult to realize: however, for the residual network, it only needs to be $f(x) = 0$, which can be converted into $H(x) = f(x) + x$ to $H(x) = x$, and the operation is relatively simple to eliminate the undesirable effect of the network layer mapping that is equivalent to directly cropping the convolutional layer.

### *B. One-dimensional residual generative adversarial networks*

#### 1) Generator network structure

In this paper, an improved one-dimensional residual generative adversarial network is proposed based on the GAN model. Firstly, the problem of network degradation is improved by introducing the residual module to improve the training stability; then, the neural network in the GAN is replaced by a one-dimensional convolutional network to make the network more compatible with the generation of one-dimensional sequences: finally, a suitable loss function is chosen to make the training more efficient and stable. The $G$ network structure of the model is shown in Figure 2.

#### 2) Structure of the discriminator network

The structure of the discriminator network ($D$ network) of 1D-RGAN is shown in Fig. 3. First, the convolutional layer is responsible for feature extraction, which uses the LeakyReLU

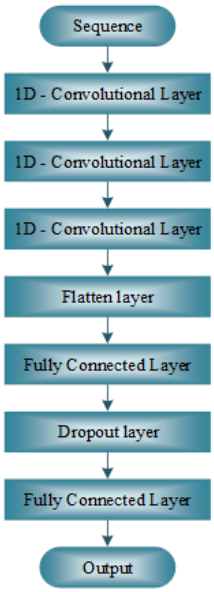


Figure 3: The discrimina tor structure of 1D-RGAN

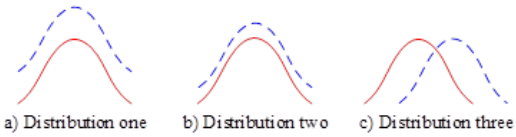


Figure 4: JS divergence diagram

activation function. Second, the Flatten layer compresses the data into one-dimensional data between the convolutional layer and the fully connected layer. Then, the Dropout layer is more effective in mitigating the occurrence of overfitting: finally, the output layer uses the tanh activation function instead of the original GAN's sigmoid activation function due to the choice of the loss function in the later section, which makes the output of the $D$-network no longer a binary classification problem between 0 and 1.

#### 3) Design of loss function

The JS scatter is illustrated schematically in Figure 4. Specifically, the dashed and solid lines depict two distinct distributions. Notably, the distance separating these distributions in Figure 4(b) is visibly closer than that in Figure 4(a). However, owing to the absence of intersection between them, this results in a constant scatter and a gradient of zero. Consequently, the loss function fails to be positively influenced by alterations in the network parameters, rendering the optimization of the network challenging. It is only when these two distributions intersect, as exemplified in Figure 4(c), that the loss function acquires significance, enabling the proper optimization of network parameters.

In this paper, Wasserstein distance ($W$ distance) is introduced as a loss function to solve this problem. The value of $W$ distance measures the distance between two distributions, as shown in (3). First, the set of all paths from $P_2$ training to

$P_t$ is obtained by calculation: then, the distance of the shortest path between $P_t$ and $P_2$, i.e., W distance, is obtained from it. Even if the two distributions do not overlap, the W distance still reflects the distance of their distributions.

$$W_{[P_t,P_p]} = \inf_{\gamma \in \Pi[P_t,P_p]} \iint \gamma(x,y)\, d(x,y)\, dxdy, \quad (3)$$

where $\prod [P_r, P_R]$ is joint distribution of $P_r$ and $P_g$, $P_r$ is true distribution, $P_g$ is generative distribution and $Inf$ is lower bound for the function value.

However, there exists a restriction on the use of the $W$ distance as a loss function for the $D$ network, namely the Lipschitz continuum. Its adds a restriction to the continuous function $f$ by requiring the existence of a constant $K \geq 0$ such that any two elements $x_1$ and $x_2$ in the domain of definition satisfy (4). In layman's terms, the Lipschitz restriction is the requirement that the function $f$ is foot-stoppingly affine in the domain of definition. In order to satisfy this condition, the authors of WGAN restrict all the parameters of the neural network $w$, to the range $[-c, c]$, when there must exist some constant $K$ such that the local variation of the function does not exceed the restriction.

$$|f(x_1) - f(x_2)| \leq K |x_1 - x_2|. \quad (4)$$

However, limiting the parameter range to $[-c, c]$ may cause problems such as vanishing gradients and gradient explosion. Only if $c$ is set appropriately, the network can be trained properly. However, it is difficult for that tuning parameter to work smoothly in practical applications. So in this paper, the gradient penalty mechanism proposed by WGAN-GP is used to solve the problem. The gradient penalty satisfies the Lipschitz continuous constraints by setting an additional loss term, and the loss function of the D network in this paper is shown in (5).

$$\begin{aligned} L(D) &= -E_{x-p_f}[D(x)] + E_{x-p_f}[D(x)] \\ &\quad +\lambda E_{x-p_1}\left[\|\nabla_x D(x)\|_p - 1\right]^2, \end{aligned} \quad (5)$$

where $\nabla$ is the sign of the gradient and $P_x$ is the distribution of random sampling points on the line connecting the real and generated data.

The gradient penalization method involves the process of mixing the true and generated samples. First, a pair of true and false samples $x_r \in P_r$, $x_2 \in P_2$ are randomly sampled, then a random number $\alpha = uniform[0, 1]$ is taken and finally sampled on the line connecting $x_r$ and $x_2$, the sampling method is shown in (6). The mixing data consists of the set of $\hat{x}$, whose distribution is $P_x$. The mixing data is fed into the mixing output (mix_output) obtained from the $D$ network, which will be involved in the computation of the loss values of the $D$ network and the $G$ network.

$$\hat{x} = \alpha x_r + (1 - \alpha) x_2. \quad (6)$$

The GP term in Eq. (5) is shown in Eq. (7). By adding the GP term and calculating the gradient of the three loss values, the gradient of the 1 network can effectively meet the requirement that the gradient of the $D$ network does not exceed $K$ so that the loss function is smooth within a certain range. In summary, in this paper, the $W$ distance with GP term will be used as the loss function of the 1-dimensional convolutional generative adversarial network.

$$\lambda E_{x \sim P_3}\left[\|\nabla_x D(x)\|_p - 1\right]^2. \quad (7)$$

### C. GAN key generation model

#### 1) Improvement of chaotic systems

In this section, an improved chaotic system is proposed for the problem of insufficient randomness of Logistic chaotic system, a new GAN key generation model is constructed, and finally, the performance and randomness of pseudo-random sequences generated by the GAN key generation model are verified, which proves that it can be applied to the field of image encryption.

Based on Logistic chaotic system, an improved quadratic chaotic system is proposed as shown in (8).

$$x_{t+1} = \text{mod}\left(a + (1 - 8x_t)^2, 1\right), (0 < x_t < 1, i = 1, 2, \ldots), \quad (8)$$

where $mod$ is mode-taking operation and $a$ is control parameters.

This chaotic system parameter $a \geq 0$, the range of parameters is increased than the Logistic chaotic system parameter range, the chaotic sequence generated by iteration is more uniformly distributed in the whole space. In order to make the chaotic system dynamics behavior more complex, the chaos interval is larger, the control parameters of the system are more, and more to meet the security needs of image encryption, this paper proposes an improved cubic chaotic system based on the Logistic chaotic system, the formula is shown in (9).

$$x_{t+1} = \sin\left(a - b x_t (1 - 8x_i)^2\right), (0 < x_t < 1, i = 1, 2, \ldots), \quad (9)$$

where control parameter $a \geq 0$, control parameter $b \in (-\infty, +\infty)$, the system has more control parameters and a wider range of chaos.

#### 2) Model structure

In this paper, we design a new key flow generation model called GAN key generation model using the 1D-RGAN and chaotic system designed above, and the structure is shown in Figure 5.

The steps of training model and key generation are as follows:

- Step 1: Generate three chaotic sequences $f_1$, $f_2$ and $f_3$ as real data under different initial conditions using the new chaotic system.
- Step 2: Input the real data to the discriminator: generate a normally distributed random noise between $[0, 1]$ to input to the generator. The generator and the discriminator are trained according to the algorithm and after the training, the generator generates random sequences $y_1$, $y_2$ and $y_3$.

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

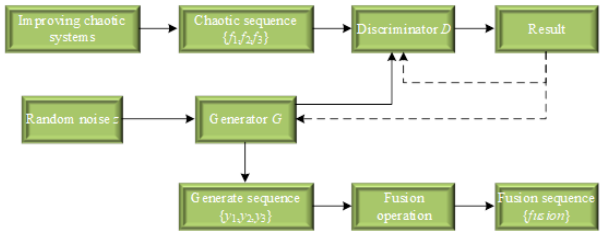


Figure 5: Structure diagram ofGAN key gencration modcl

- Step 3: Fuse $y_1$, $y_2$ and $y_3$ to generate the fused random sequence $fusion\,(i)\,i = 1, 2, \ldots, M \times N$. $fusion\,(i)$ is computed as follows, first quantize the values of the three sequences into the interval [0, 255] to obtain the decimal quantized sequences $s_1$, $s_2$ and $s_3$.

$$s_1(i) = \text{mod}\,(y_1(i)y_2(i) - y_3(i), 256)\,, \tag{10}$$

$$s_2\,(i) = \text{mod}\left(y_1\,(i) + s_1^2\,(i)\,, 256\right), \tag{11}$$

$$s_3\,(i) = \text{mod}\left(y_2\,(i) + s_2^2\,(i)\,, 256\right). \tag{12}$$

Then, in order to sufficiently disrupt the correlation between the sequences, the intermediate sequences $middle$, $m_1$ and $m_2$ are computed. when $i = 1$, the first value of the sequence is computed.

$$\text{middle}\,(1) = \text{mod}\,(s_1\,(1)\,s_2\,(1) - s_3\,(1)\,, 256)\,, \tag{13}$$

$$m_1\,(1) = \text{mod}\left(\text{middle}\,(1)^2 + s_1\,(1)\,, 256\right), \tag{14}$$

$$m_2\,(1) = \text{mod}\left((m_1\,(1))^2 + s_2\,(1)\,, 256\right). \tag{15}$$

When $1 < i \leq M \times N$, the intermediate sequences $middle\,(i)$, $m_1\,(i)$, and $m_2\,(i)$ are computed, correlating the value in bit $i$ with the value in bit $i-1$.

$$\text{middle}\,(i) = \text{mod}\,(s_1\,(i)\,s_2\,(i) - \text{middle}\,(i-1)\,, 256)\,, \tag{16}$$

$$m_1\,(i) = \text{mod}\left(\text{middle}\,(i)^2 + s_1\,(i)\,, 256\right), \tag{17}$$

$$m_2\,(i) = \text{mod}\left(m_1\,(i)^2 + s_2\,(i)\,, 256\right). \tag{18}$$

Finally, the fusion sequence $\text{fusion}\,(i)$ is calculated as shown in (19).

$$\text{fusion}\,(i) = \text{mod}\,(m_1\,(i) + m_2\,(i)\,, 256)\,i = 1, 2, \ldots, M \times N. \tag{19}$$

- Step 4: Save the model parameters so that the generator for the GAN key generation model can be invoked directly.

## III. Generative Adversarial Based Image Encryption Algorithm

### A. The knuth-durstenfeld shuffling algorithm and its improvements

The Knuth-Durstenfeld algorithm stands as one of the seminal shuffling algorithms, whose concept is rooted in card drawing, swapping, and insertion mechanisms. Specifically, the Fisher-Yates Shuffle algorithm aligns with the card drawing paradigm, aiming to randomly extract data from the original data set without repetition, thereby introducing a certain level of randomness. Conversely, the Knuth-Durstenfeld algorithm, akin to card swapping, refines the Fisher-Yates approach by initiating the traversal from the tail of the array. In this method, an unprocessed element is randomly exchanged with the current tail element of the array, whereupon the tail advances, and the process repeats until all elements have been traversed. Notably, this algorithm eschews the use of an auxiliary array for data storage, contributing to its reduced space and time complexity.

However, the direct application of the Knuth-Durstenfeld shuffling algorithm in image encryption encounters limitations. The primary obstacles are twofold: firstly, the original algorithm's random number generator is incapable of reproducing the exact sequence during decryption, thus hindering successful decryption, as it fails to mirror the encryption process's randomness. Secondly, as the algorithm progresses, the unprocessed data in the array diminishes, leading to a corresponding decrease in the random number generator's range. Notably, both chaotic sequences are confined to a fixed range and do not diminish with iteration. To address these issues, this paper proposes an enhancement to the Knuth-Durstenfeld algorithm by substituting the random number generator with a sequence generated by a chaotic system. Furthermore, the implementation incorporates a modulo function to ensure that the generated random numbers remain within the bounds of the unprocessed data array, thereby overcoming the aforementioned limitations.

### B. Overall framework of image encryption algorithm

The overall flowchart of the proposed Generative Adversarial Network (GAN)-based image encryption algorithm incorporating plaintext correlation is presented in Figure 6. This algorithm is systematically divided into two distinct phases. The initial phase pertains to key generation, where the key generation network undergoes training to establish a mapping from the original image to a key style image. Notably, the key style image is chosen as a highly random, uniformly distributed image, enabling the network to generate a corresponding key image tailored to the input original image. This generated key is inherently linked to the plaintext, thereby enhancing the security level of the associated encryption algorithm. Within the framework of this paper's encryption algorithm, the initial step involves inputting the original image into the key generator to retrieve the corresponding key.The subsequent phase constitutes the encryption stage, where the

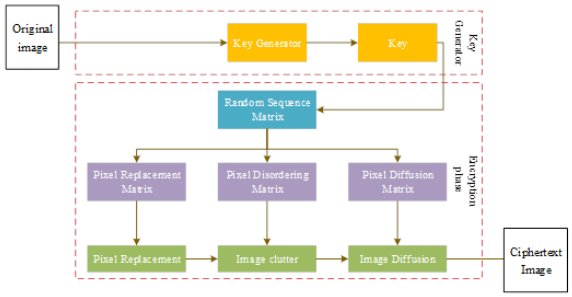


Figure 6: Generate adversarial network and plaintext related image encryption algorithm

generated key image is transformed into a matrix representation of a random sequence. Subsequently, the encryption of the original image is accomplished through a series of operations: pixel value substitution, image disambiguation, and image diffusion, ultimately yielding the ciphertext image. The decryption process is simply the reverse of the encryption procedure, underscoring the symmetric nature of the overall encryption algorithm. Consequently, the decrypted image remains identical to the original image, ensuring data integrity throughout the encryption-decryption cycle.

### C. Image encryption

The encryption algorithm presented in this paper leverages the key image, generated from the input original image during the key generation stage, and subsequently transformed into a matrix of random numbers. This matrix then facilitates the encryption of the image through a comprehensive process comprising three distinct steps: pixel value substitution, image disambiguation, and image diffusion. This orchestrated sequence of operations culminates in the production of a high-security ciphertext image.

#### 1) Pixel value replacement process

The first stage of the encryption scheme is the pixel value replacement, corresponding to the key image generated by the key generation network, which is converted into three pixel value replacement matrices based on the RGB three-channel information of the key image $PR_{mat}$ . Through the pixel value replacement formula, the operation is performed on the three channels corresponding to the plaintext image, and a new pixel value is obtained to replace the original pixel value of the plaintext image, which completes the first encryption step and obtains the replacement image. The pixel replacement formula is shown in (20).

$$\text{NewValue} = (\text{ImageList}[\text{index}] + \text{KeyList}[\text{index}]) \,\%\, (256), \tag{20}$$

where ImageList is the pixel matrix of the original image, index is the corresponding coordinate value, KeyList is the pixel replacement matrix to which the key image is converted, and New Value is the new pixel value at the current position

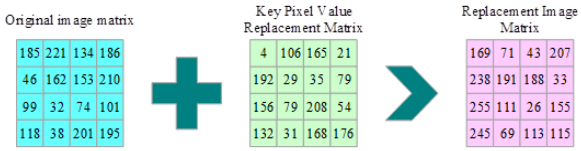


Figure 7: Pixel value replacement process on 4 ×4 pixel matrix

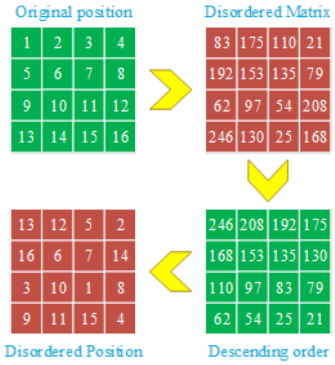


Figure 8: Image scrambling process on 4× 4 pixel matrix

after replacement. Figure 7 demonstrates how the pixel value replacement process operates on a 4×4 pixel matrix.

#### 2) Image disordering process

The key image generated by the key generation network is converted into three disordered matrices $PR_{mat}$ in the image disarray stage, indexed with the replacement image after pixel value replacement is completed in the previous stage, and the position information in the replacement image corresponds one-to-one with the position in the disordered matrices, and the values of the three disordered matrices $PR_{mat}$ are rearranged in accordance with the ascending or descending algorithm in the image disarray stage, and the corresponding replacement image follows the rearranged matrix to disrupt the original position, and the replacement image completes the image disarray process with the ordering of the disordered matrix, obtaining the disarrayed image. The corresponding replacement images will follow the rearrangement of the disordered matrices to disrupt the original positions, and the replacement images are sorted with the disordered matrices to complete the image disruption process and obtain the disrupted image. In Fig. 8, it is demonstrated how the image disarrangement algorithm can be accomplished by using the descending order within a 4×4 pixel matrix.

#### 3) Image diffusion process

In the algorithm of this chapter, the scrambled image obtained after the image scrambling stage is converted into binary form by converting the pixel values in each plane of the RGB channel, and the pixel values of the three random number matrices $PR_{mat}$ of the key image are also converted into binary form, which are then used in the different-or-ortho (XOR)

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

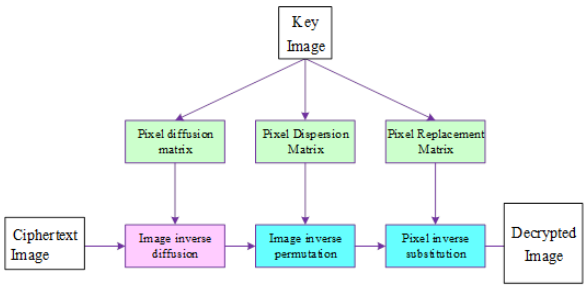


Figure 9: Decryption algorithm flow chart

operation with the binary pixel values of each corresponding position in the scrambled image to get the new pixel values, which are then assigned to the corresponding positions and the three matrices after the operation are recombined into a three-channel image to get the final ciphertext image. The three matrices after the operation are recombined into a three-channel image, and the final ciphertext image is obtained. The encryption algorithm designed in this chapter realizes the diffusion phase of the encrypted image through the different-or operation, which enhances the security of the ciphertext image.

### D. Image decryption

The encryption algorithm introduced in this paper adheres to a symmetric encryption paradigm, where the decryption process serves as the inverse mirror of its encryption counterpart, necessitating possession of the encryption key by the decrypting party. During encryption, the algorithm sequentially executes three primary operations: pixel value substitution, followed by disorganization of the substituted results, and concluding with diffusion on the disorganized matrix. Consequently, the decryption process reverses this order, as illustrated in Figure 9. Initially, an inverse diffusion operation is applied to the ciphertext image, subsequently, the disrupted order is reinstated using the same sorting algorithm employed during disorganization, and finally, an inverse pixel value replacement is executed to yield the decrypted image. Given the reversibility of the proposed encryption scheme, the decrypted image remains identical to the original plaintext image.

## IV. Findings and Analysis

In this experimental setup, we employed standard grayscale images of Lena, Peppers, Baboon, and Airplane, each with dimensions of 512×512 pixels, as our test subjects. The outcomes of the encryption and subsequent decryption processes applied to these images are presented in Figure 10. Upon inspection, it becomes evident that the visual fidelity between the original plaintext image and its decrypted counterpart is impeccable, demonstrating a perfect match. Furthermore, the encrypted images resemble a chaotic noise pattern, devoid of any discernible or effective information leakage, thereby validating the security of the encryption method.

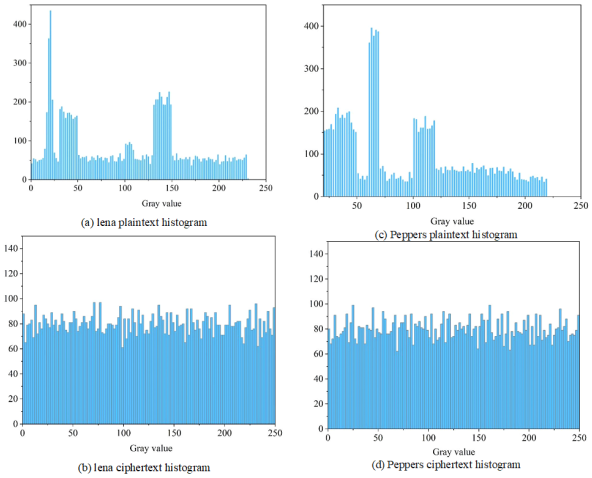


Figure 10: Lena and Peppers plaintext and ciphertext histograms

### A. Histogram analysis

Histogram serves as a prevalent criterion for assessing the experimental outcomes of image encryption algorithms. For an encryption algorithm to demonstrate superior performance, the histogram of its resulting ciphertext image ought to exhibit uniformity. In this study, the images of Lena and Peppers have been chosen as the experimental subjects for investigation.

The analysis of the histograms for the plaintext and ciphertext of Lena and Peppers images is presented in Figure 10. Inspection of the figure reveals that the histograms of the plaintext images exhibit irregular statistical distributions. Specifically, Lena's plaintext image displays a peak in statistical frequency within the gray value interval of 15-25, while Peppers' plaintext image exhibits a similar peak within the range of 65-70. In contrast, the histogram of the ciphertext image reveals a stable and flat gray value distribution, devoid of significant fluctuations. These experimental outcomes demonstrate that the proposed algorithm effectively obscures the statistical properties of the gray values in the test images, thereby exhibiting robust diffusion characteristics and resilience against statistical attacks.

### B. Neighboring pixel correlation analysis

The original image exhibits a considerable degree of correlation among neighboring pixels. When this correlation is pronounced, an attacker may readily extract image information through statistical analysis. Conversely, as the correlation coefficient between pixels approaches zero, statistical analysis attacks become increasingly challenging, thereby enhancing the security of the encryption algorithm.

The outcomes of the correlation analysis experiment are detailed in Table 1, which compares the correlation coefficients of grayscale images against a benchmark. A correlation coefficient value approaching 1 signifies a strong correlation among pixels, whereas values nearing 0 indicate a weaker correlation. Notably, the experimental ciphertext images of Lena

and Peppers exhibit correlation coefficients highly proximal to the ideal value of 0, in contrast to the plaintext images, whose correlation coefficients tend towards 1. This observation attests to the effective scrambling of the experimental ciphertext images, where pixels have been thoroughly rearranged and dispersed, rendering the image information more resilient to analysis and decryption. Consequently, the experimental results validate that the proposed scheme significantly reduces the correlation among image pixels, thereby demonstrating a commendable encryption efficacy.

### C. Robustness analysis

In order to verify the robustness of the algorithms in this chapter, this chapter analyzes the proposed algorithms from two aspects: noise attack and cropping attack.

The robustness of the proposed algorithm is analyzed from two aspects.

#### 1) Noise Attack Analysis

Embedding the ciphertext within a carrier image offers visual obscurity during transmission and the associated encryption key serves as a deterrent against attacks. Nevertheless, the visually secure image may be subjected to various assaults during transmission, including noise contamination and data loss, which can complicate the process of image decryption. Consequently, a thorough analysis of the algorithm's resilience against such attacks is imperative.

The peak signal-to-noise ratio (PSNR) values under noise attack are shown in Table 2. It can be seen that whether it is Gaussian noise or Pepper noise, the PSNR value decreases with the increase in the degree of added noise. When the degree of noise is 0.0002, the PSNR value is the largest, which is 36.1592 under Gaussian noise and 36.5694 under pretzel noise. moreover, the PSNR values of the four images do not differ much under various degrees of noise attacks, which indicates that each image is affected by the noise attack in a more average way, and each steganographic image is affected by the basically the same.

#### 2) Cropping Attack Analysis

To test the robustness of cropping attack, the ciphertext image is cropped in different sizes and decrypted in the same steps. The Lena image is used as the test image. To further evaluate the quality of decrypted image after cropping attack, the values of PSNR and SSIM of plaintext image and recovered image can be calculated to measure and evaluate the difference between different decrypted images.

The values of SSIM and PSNR against cropping and pepper noise are shown in Table 3. From the table, it can be seen that all the values of PSNR are greater than 40dB and all the values of SSIM are close to 1, which proves that the proposed algorithm can effectively resist noise and cropping attacks.

### D. Analysis of anti-differential attacks

Optimal encryption algorithms ought to exhibit robust resistance against differential attacks, characterized by a marked disparity in encryption efficacy when subtle alterations are introduced to the plaintext image utilized for encryption. This feature hinders attackers from extracting meaningful information from the encrypted image. Two key metrics commonly employed to assess an algorithm's resilience against differential attacks are the Number of Pixels Change Rate (NPCR) and the Uniform Average Change Intensity (UACI). Specifically, NPCR quantifies the percentage of pixel alterations in the corresponding ciphertext image resulting from a random change in a single pixel's value within the plaintext image. Meanwhile, UACI measures the intensity variation in the pixel values of the corresponding ciphertext image when a single pixel's value undergoes a change.

The results of the simulation experiments pertaining to the Number of Pixels Change Rate (NPCR) and Uniform Average Change Intensity (UACI) are presented in Table 4, alongside a comparative analysis. As evident from the table, the NPCR simulation outcomes achieved by the algorithm proposed in this paper approximate the ideal value of 1, surpassing those of other algorithms. This indicates a superior competitive edge, thereby attesting to the algorithm's robust performance against differential attacks and its high security prowess.

GAN generates a model of chaotic sequences that can be used to reconstruct the chaotic sequence of the generated chaos, but the attacker must know that GAN is competing against the parameters, the initial state of the generated network and the arrangement of the matrix. At the same time, the attacker can use the method of poor attack to guess the key, but this method will produce high cost in calculation. Of course, the attacker can also use other methods such as plain attack, difference attack and linear attack method to attack the model, and if you use these three methods, you must solve the network's parameters, the initial values and the input, which is extremely difficult in the case of the secret text.

## V. Conclusion

This paper presents a novel image encryption algorithm, integrating the GAN-based key generation model with the Knuth-Durstenfeld shuffling algorithm. The subsequent conclusions are drawn based on a comprehensive analysis of the following aspects: the histogram distribution of the ciphertext image, the correlation among neighboring pixels, the algorithm's robustness, and its resistance to differential attacks.

1) The histogram of the plaintext image exhibits irregular statistical characteristics, with Lena's plaintext image displaying the highest frequency of gray values within the range of 15-25, whereas Peppers' plaintext image exhibits the peak frequency in the range of 65-70. Conversely, the gray value distribution of the ciphertext image's histogram remains relatively stable and smooth, devoid of significant fluctuations.
2) The correlation coefficients of the experimental ciphertext images, Lena and Peppers, closely approximate the ideal value of 0, whereas the correlation coefficients of the corresponding plaintext images tend towards 1. This validates that the experimental ciphertext images

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

| | | Lena | Peppers | All black | All white |
|---|---|---|---|---|---|
| Plaintext image | Horizontal direction | 0.9465 | 0.9814 | NaN | NaN |
| | Vertical direction | 0.9610 | 0.9598 | NaN | NaN |
| | Diagonal direction | 0.9016 | 0.9238 | NaN | NaN |
| Ciphertext image | Horizontal direction | -0.0051 | 0.0053 | -0.0072 | -0.0068 |
| | Vertical direction | -0.0058 | 0.0051 | -0.0134 | -0.0023 |
| | Diagonal direction | -0.0014 | 0.0018 | 0.0042 | -0.0018 |

Table 1: Correlation experiment result

| Attack | | Gaussian noise | | | | Salt-and-pepper noise | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variance/noise density | | 0.0002 | 0.0004 | 0.0006 | 0.0008 | 0.0002 | 0.0004 | 0.0006 | 0.0008 |
| This algorithm | Lena | 36.1592 | 35.9876 | 35.1258 | 35.0921 | 36.5694 | 35.8714 | 35.2164 | 34.4164 |
| | Peppers | 37.2413 | 36.5496 | 36.3984 | 36.1891 | 37.0130 | 37.4851 | 36.1456 | 35.9741 |
| | Baboon | 34.4896 | 34.2897 | 34.1651 | 34.0236 | 34.3684 | 33.8974 | 33.0478 | 32.1026 |
| | Airplane | 34.0124 | 33.9874 | 33.6875 | 33.1052 | 35.2189 | 35.0127 | 34.6587 | 34.3697 |

Table 2: Comparison of anti-noise attack capability of image hiding algorithms

| Attack mode | Evaluation index | Lena | Peppers | Baboon | Airplane |
|---|---|---|---|---|---|
| Data loss | PSNR | 46.1254 | 45.8654 | 44.9743 | 47.8745 |
| (64×64) | SSIM | 0.7941 | 0.8436 | 0.9784 | 0.7746 |
| Noise attack | PSNR | 47.6845 | 46.4789 | 45.9648 | 43.9876 |
| (0.01) | SSIM | 0.8761 | 0.8974 | 0.9741 | 0.8111 |

Table 3: SSIM and PSNR values of anti-clipping and salt-and-pepper noise

| | Encrypted image | Textual algorithm | 3D image encryption algorithm | PSO-BP neural network algorithm | DNA cross mutation encryption algorithm | Piecewise chaotic mapping algorithm |
|---|---|---|---|---|---|---|
| NPCR | Lena | 0.9986. | 0.9978 | 0.9912 | 1.0461 | 0.9949 |
| | Peppers | 0.9967 | 0.9605 | 0.9935 | 0.9953 | 0.9925 |
| | Camera | 0.9961 | 0.9871 | 0.9967 | 0.9345 | 0.9934 |
| UACI | Lena | 0.4398 | 0.4056 | 0.3974 | 0.4196 | 0.4311 |
| | Peppers | 0.4257 | 0.4185 | 0.4186 | 0.4177 | 0.4209 |
| | Camera | 0.4305 | 0.4113 | 0.4165 | 0.4231 | 0.4298 |

Table 4: Simulation results of NPCR and UACI

undergo effective scrambling, with their pixels being thoroughly rearranged and dispersed. Consequently, the information contained within these images becomes more challenging to analyze and decrypt. This underscores the algorithm's efficacy in diminishing pixel correlation and demonstrates its impressive encryption performance.

3) The SSIM (Structural Similarity Index Measure) and PSNR (Peak Signal-to-Noise Ratio) values, evaluated against cropping and pepper noise attacks, consistently exceed 40dB, with SSIM values nearing unity. This robust performance substantiates the algorithm's capability to effectively withstand noise and cropping attacks, highlighting its resilience and suitability for secure image transmission and storage.
4) The simulation experimental results of the Number of Pixels Change Rate (NPCR) obtained using the algorithm presented in this paper are proximal to the ideal value of 1, consistently surpassing those achieved by other algorithms. This competitive edge underscores the algorithm's exceptional performance in resisting differential attacks, thereby contributing to its heightened security capabilities.

Although the model of the paper has been preliminarily implemented, in the future research, there is a lot of experiments, the evaluation of the sorting parameters, and the ability to continue to improve the performance of the encryption method.

## Acknowledgments

This work was supported by the Science and Technology Project of Henan (Optical image compression and encryption based on depth learning and high-quality reconstruction, 242102210058).

## References

[1] He, Q., & He, H. (2020). A novel method to enhance sustainable systems security in cloud computing based on the combination of encryption and data mining. *Sustainability, 13*(1), 101.
[2] Ogiela, M. R. (2019). Cognitive solutions for security and cryptography. *Cognitive Systems Research, 55*, 258-261.
[3] Qiu, J. (2022). Ciphertext database audit technology under searchable encryption algorithm and blockchain technology. *Journal of Global Information Management (JGIM), 30*(11), 1-17.
[4] Li, Q., Ju, Y., Zhao, C., & He, X. (2021). An encrypted field locating algorithm for private protocol data based on data reconstruction and moment eigenvector. *IEEE Access, 9*, 42947-42958.
[5] Zhang, X. J., Li, Z., & Deng, H. (2017). Information security behaviors of smartphone users in China: an empirical analysis. *The Electronic Library, 35*(6), 1177-1190.
[6] Ri, Y., & Fujimoto, H. (2019). Practical Algorithm Design for the FFT-based Robust Image Registration Method. *IEEJ Transactions on Industry Applications, 139*(1), 22-29.
[7] Wu, Y., Feng, G., & Fung, R. Y. (2018). Comparison of information security decisions under different security and business environments. *Journal of the Operational Research Society, 69*(5), 747-761.
[8] Ehsaeyan, E., & Zolghadrasli, A. (2021). A Darwinian differential evolution algorithm for multilevel image thresholding. *International Journal of Humanoid Robotics, 18*(04), 2150012.
[9] Shaukat, K., & Hassan, M. U. (2017). Cloud security through encryption techniques. *Transylvanian Rev, 1*.

[10] Huang, W., Jiang, D., An, Y., Liu, L., & Wang, X. (2021). A novel double-image encryption algorithm based on Rossler hyperchaotic system and compressive sensing. *IEEE Access, 9*, 41704-41716.
[11] Wang, N., Jiao, J., Zhang, S., & Fu, J. (2022). Secure and Efficient Semantic Extension Search over Encrypted Documents by Integrating Cloud and Fog Systems. *Mathematical Problems in Engineering, 2022*(1), 9349651.
[12] Catak, F. O., Aydin, I., Elezaj, O., & Yildirim-Yayilgan, S. (2020). Practical implementation of privacy preserving clustering methods using a partially homomorphic encryption algorithm. *Electronics, 9*(2), 229.
[13] Jia, S., Ma, L., & Qin, D. (2018). Research on scene understanding-based encrypted image retrieval algorithm. *IEEE access, 7*, 6587-6596.
[14] Nguyen, T. S., Chang, C. C., & Lin, C. C. (2022). High capacity reversible data hiding scheme based on AMBTC for encrypted images. *Journal of Internet Technology, 23*(2), 255-266.
[15] Tang, Z., Pang, M., Yu, C., Fan, G., & Zhang, X. (2021). Reversible data hiding for encrypted image based on adaptive prediction error coding. *IET Image Processing, 15*(11), 2643-2655.
[16] Ye, G., Pan, C., Dong, Y., Jiao, K., & Huang, X. (2021). A novel multi-image visually meaningful encryption algorithm based on compressive sensing and Schur decomposition. *Transactions on Emerging Telecommunications Technologies, 32*(2), e4071.
[17] Li, X., Zhou, C., & Xu, N. (2018). A secure and efficient image encryption algorithm based on dna coding and spatiotemporal chaos. *Int. J. Netw. Secur., 20*(1), 110-120.
[18] Hu, W. W., Zhou, R. G., Luo, J., Jiang, S. X., & Luo, G. F. (2020). Quantum image encryption algorithm based on Arnold scrambling and wavelet transforms. *Quantum Information Processing, 19*, 1-29.
[19] Rahman, Z. A. S., Jasim, B. H., Al-Yasir, Y. I., & Abd-Alhameed, R. A. (2022). Efficient colour image encryption algorithm using a new fractional-order memcapacitive hyperchaotic system. *Electronics, 11*(9), 1505.
[20] Farwa, S., Muhammad, N., Shah, T., & Ahmad, S. (2017). A novel image encryption based on algebraic S-box and Arnold transform. *3D Research, 8*, 1-14.
[21] Kumar, P., & Diwakar, M. (2021). A novel approach for multimodality medical image fusion over secure environment. *Transactions on Emerging Telecommunications Technologies, 32*(2), e3985.