sensors

MDPI

Article

Face Image Encryption Based on Feature with Optimization Using Secure Crypto General Adversarial Neural Network and Optical Chaotic Map

Majed Alsafyani [1], Fahad Alhomayani [2], Hatim Alsuwat [3] and Emad Alsuwat [1,*]

[1] Department of Computer Science, College of Computers and Information Technology, Taif University, Taif 26571, Saudi Arabia
[2] Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif 26571, Saudi Arabia
[3] Department of Computer Science, College of Computer and Information Systems, Umm Al-Qura University, Makkah 24382, Saudi Arabia
[*] Correspondence: alsuwat@tu.edu.sa

Abstract: Demand for data security is increasing as information technology advances. Encryption technology based on biometrics has advanced significantly to meet more convenient and secure needs. Because of the stability of face traits and the difficulty of counterfeiting, the iris method has become an essential research object in data security research. This study proposes a revolutionary face feature encryption technique that combines picture optimization with cryptography and deep learning (DL) architectures. To improve the security of the key, an optical chaotic map is employed to manage the initial standards of the 5D conservative chaotic method. A safe Crypto General Adversarial neural network and chaotic optical map are provided to finish the course of encrypting and decrypting facial images. The target field is used as a "hidden factor" in the machine learning (ML) method in the encryption method. An encrypted image is recovered to a unique image using a modernization network to achieve picture decryption. A region-of-interest (ROI) network is provided to extract involved items from encrypted images to make data mining easier in a privacy-protected setting. This study's findings reveal that the recommended implementation provides significantly improved security without sacrificing image quality. Experimental results show that the proposed model outperforms the existing models in terms of PSNR of 92%, RMSE of 85%, SSIM of 68%, MAP of 52%, and encryption speed of 88%.

Keywords: information security; biometrics; face feature encryption; image optimization; cryptography; deep learning

check for updates

Citation: Alsafyani, M.; Alhomayani, F.; Alsuwat, H.; Alsuwat, E. Face Image Encryption Based on Feature with Optimization Using Secure Crypto General Adversarial Neural Network and Optical Chaotic Map. *Sensors* **2023**, *23*, 1415. https://doi.org/10.3390/s23031415

Academic Editors: Piyush Kumar Shukla, Ahmed A. Abd El-Latif, Rupak Kharel, Udai Pratap Rao and Prashant Kumar Shukla

Received: 13 December 2022
Revised: 19 January 2023
Accepted: 20 January 2023
Published: 27 January 2023

1. Introduction

DL is a very powerful technique in computer vision applications. Surveillance is a high-potential application field for DL. DL requires a large training data set to achieve excellent performance. However, collecting enough training data while maintaining the anonymity of people in data is expensive, especially for surveillance applications. No one wants their images to be included in the dataset because developers can monitor anyone's actions. A comparable circumstance can be found during a surveillance operation. A security camera owner can monitor anything. Image encryption is one of the methods for maintaining privacy [1]. The picture encryption process converts an original image into an encrypted image in which no one can recognize the contents. Image encryption methods were primarily created to send photos securely over a public network. For people and machines to recognize the contents of an encrypted image, the image must first be decrypted. Anyone may recognize the contents of the image once it has been decrypted. This means that picture decryption can compromise privacy [2].

People have entered the era of big data as a result of technological advancements. Internet and computer technology have advanced swiftly, network popularity has grown significantly, and information interaction technology has matured. While most people utilize the Internet to send information, it also creates many data security risks. Data transmission security is rapidly affecting the security of individuals, businesses, and even countries as networks expand into new domains. Because of its visual features, an image has a strong expression influence on the data it contains. Many information expressions favor visuals because of their widespread use in information interaction. Owners of valuable photographs frequently utilize the Internet to conduct an auction or post their image data. That strategy removes geographical limits such as geography, and it is convenient and quick, but it also saves money [3]. However, throughout the network transmission process, insecure elements of picture data provide an opportunity for malicious attacks, and original image data may be attacked, resulting in data leakage or destruction. The goal of the picture encryption technique is to enhance the security of image data, minimize the risk of data leakage and destruction, and ensure the secure transmission of original data. In a few cases, image data are encrypted before transmission. For example, before medical images may be transmitted over the Internet, they must be encrypted to safeguard patient privacy. Criminal attacks such as destructive damage and data theft are common, and picture encryption technology is continually evolving. An urgent topic to be tackled is how to increase picture security, key transmission security, and anti-attack capability [4]. The iris, face, fingerprint, voice, deoxyribonucleic acid (DNA), and palm positions are examples of information properties created by human tissue structures. Because of their uniqueness, biological traits of the human body are extensively utilized to determine recognition and other sectors. The properties of the iris are extracted using iris recognition technology. It is part of the human biological feature extraction technology and is of extremely high grade. It is better for picture encryption because it improves the algorithm's security and anti-attack capability. The identity recognition method based on iris feature extraction is receiving increased attention in academic and industry sectors. It has an extensive variety of applications and is progressively implemented in various departments with high security needs, such as finance and secrecy [5]. Applying DL techniques to the area of image security to resolve classic challenges has also received much attention recently and has made significant progress. However, many researchers are interested in how to better utilize the benefits of DL in image cryptography, image authentication, and image steganography. To assist relevant researchers in better understanding the field of DL uses in digital image security together with its upcoming progress, the origin and development method of DL techniques in image cryptography, steganography, and authentication were organized from numerous perspectives in this paper, as shown in Figure 1. We then evaluate these strategies, assess their benefits and drawbacks, and make recommendations for future research on this subject.

Various security techniques are currently available to assist in repelling picture-based attacks, but they are not effective in balancing security and image quality demands. Aside from that, chaotic map behavior provides a high level of security. As a result, combining DL and chaotic behavior can provide a superior picture encryption solution. As a result, the suggested article proposes a model in which a DL chaotic map is employed to perform better optimization to improve picture encryption performance. In computer vision applications, DL is a very powerful technology. Surveillance is a high-potential application field for DL. DL requires a large training data set to achieve excellent performance. However, collecting enough training data while maintaining the anonymity of people in the data is expensive, especially for surveillance applications. No one wants their images to be included in the collection because developers can monitor anyone's actions in real-time. During a surveillance operation, a similar predicament can be encountered. A security camera owner can monitor anything. Image encryption is one of the methods for maintaining privacy. Image encryption converts an original image into an encrypted image in which the contents of the original image are unrecognizable. Image encryption methods were created primarily for the purpose of securely transmitting photos over a public network. For people and/or machines to recognize the contents of an encrypted image, the image must first be decrypted. Anyone may recognize the contents of the image once it has been decrypted. This means that image decryption has the potential to infringe on one's privacy [6].

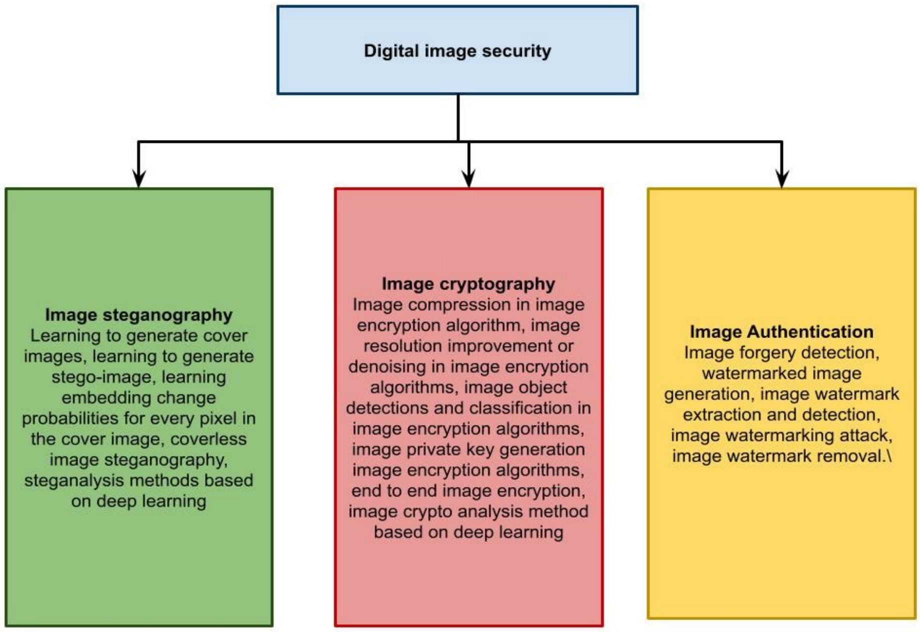

Digital image security
Image steganography
Learning to generate cover images, learning to generate stego-image, learning embedding change probabilities for every pixel in the cover image, coverless image steganography, steganalysis methods based on deep learning
Image cryptography
Image compression in image encryption algorithm, image resolution improvement or denoising in image encryption algorithms, image object detections and classification in image encryption algorithms, image private key generation image encryption algorithms, end to end image encryption, image crypto analysis method based on deep learning
Image Authentication
Image forgery detection, watermarked image generation, image watermark extraction and detection, image watermarking attack, image watermark removal.\

Figure 1. The entire architecture of the digital image security survey.

The contributions of this research are as follows:

- To propose a novel technique for face feature encryption with image optimization using cryptography and deep learning architectures;
- To develop a secure Crypto General Adversarial neural network and optical chaotic map for encryption and decryption of face images with optimization of images.

The rest of this research article is organized as follows. Section 2 of this paper shows the related work. In Section 3, novel techniques for face feature encryption with image optimization using cryptography and deep learning structural design are deliberated. In Section 4, experimental analysis and discussions are displayed. Section 5 of this paper contains its conclusion.

2. Related Work

This section contains traditional study projects that are only focused on image security. Furthermore, the literature uses picture encryption using chaotic methods [7]. However, achieving a suitable balance between security efficiency and encryption significance is difficult with these methods. DL, which uses multilayer neural networks (NNs) to extract features from raw input photos, has also attracted much interest in solving the problem. The advantages of convolution neural networks (CNNs) [8] are established in computer vision applications and picture domain transfer [9]. Image transfer from one domain to another is thought of as a texture transfer issue, to learn mapping connection amid an input image and output image from a set of matched image pairs. The most common image-to-image conversion approach is the cycle-reliable adversarial system [5], which offers two-cycle reliability losses that shift the image from one domain to another and then rebuilds back to the original image. DL technique is used to manage the image-denoising problem [10]. Image noise is interference data in image data that cause some useful image data to become invisible. The process of image denoising is considered image restoration [11]. The authors of [12] proposed a new cross-image, pixel scrambling-based rotation domain, dual-image encryption technique. In [13], the authors suggested a dual-image encryption method that incorporates DNA spatiotemporal chaos, deletion, and insertion to improve the security of the encryption process, and scramble real as well as imaginary sections of data produced by every round of encryption [14]. The approach encrypts two images simultaneously by combining DNA sequence insertion and deletion methods with scrambling and diffusion methods. Reference [15] introduced a new dual-image compression encryption method that improves the secrecy and resilience of the dual-image encryption approach. These techniques can employ the same encryption method to encrypt two photos separately, which requires decryption twice to extract the two images [16]. The authors of [17] proposed a method for medical picture cryptography based on a combination of chaotic and neural networks in their study. The major goal of the proposed method is to verify the safety of medical photographs using a less sophisticated method than current methods. Test findings supported the proposed method's performance and efficiency, which meets digital imaging and communications in medicine criteria. In [18], the authors proposed a new multikey compressed sensing and ML privacy-preserving computing system. A user, a cloud, and a trusted third party make up this computer architecture, and the trusted third party is in charge of distributing random compressed sensing keys. In [19], a machine learning technique was used for an issue involving health, commercial, or other sorts of sensitive data, which necessitates not only precise estimates. Because the cloud does not have access to keys required to decrypt the data, encryption assures that it remains private. Table 1 presents a summary of image encryption techniques, including their benefits and limitations.

Table 1. Different image encryption techniques.

Sr. No.	Image Encryption Technique	Overview	Advantages	Disadvantages
1.	Image encryption based on a public key [20]	Public key encryption uses a pair of keys, one for encryption and one for decryption, providing secure communication and nonrepudiation for image data.	Public key encryption provides secure communication as only the intended recipient can decrypt the image using their private key. It also allows for nonrepudiation and the ability to encrypt large amounts of data.	Public key encryption can be slower and more computationally expensive than symmetric key encryption. Additionally, managing and securely distributing the public and private keys can be complex and difficult.
2.	Chaos-based encryption technique [21–23]	Random starting circumstances. Numerous iterations are required; a sophisticated mapping process.	Chaos-based encryption techniques use chaotic systems to generate encryption keys, providing high levels of security and randomness. They also have the ability to resist known plaintext attacks and are resistant to differential cryptanalysis.	Chaos-based encryption techniques can be complex to implement and may have limitations in terms of encryption speed and scalability. They also may be sensitive to initial conditions and perturbations in the chaotic system.
3.	Visually meaningful image encryption technique [24,25]	Mentions an image that is at least twice as large as the original. A successful embedding method. A powerful encryption method.	Visually meaningful image encryption techniques help to preserve the visual features of an image while still encrypting it, making it more user-friendly and easy to understand. This also allows for more efficient and effective image transmission and storage.	Visually meaningful image encryption techniques may not provide as much security as other encryption methods and can be vulnerable to attacks such as stegonography and visual cryptanalysis. Additionally, it may be more computationally expensive and complex to implement.
4.	Partial image encryption techniques [26,27]	Extraction of important areas from images. Any safe encryption method.	Partial image encryption techniques allow for selective encryption of important or sensitive parts of an image, enhancing security while preserving the overall visual quality of the image. It also allows for more efficient storage and transmission as only certain parts of the image are encrypted.	Partial image encryption techniques may not provide as much security as full image encryption, as attackers may focus on the unencrypted parts of the image. It also may be more complex to implement and may require additional information to properly decrypt the image.

Table 1. *Cont.*

Sr. No.	Image Encryption Technique	Overview	Advantages	Disadvantages
5.	Symmetric key encryption techniques [28–30]	Symmetric key confidentiality. Mechanism for safe key sharing and codec conformity.	Symmetric key encryption uses the same key for encryption and decryption, providing fast and efficient encryption. It also requires less computational power and is simpler to implement compared to other encryption methods, making it more practical for many use cases.	Symmetric key encryption requires secure key distribution and management, as the same key is used for encryption and decryption, if the key is compromised the security of the encrypted data is lost. It also does not provide nonrepudiation, meaning that the sender and receiver cannot prove who sent the message.
6.	Proposed encryption technique based on cryptography and deep learning (DL) architectures.	Here, the input face image is processed and mapped using optical chaotic maps, which are utilized for efficient encryption and decryption of the image.	The proposed encryption technique provides high security by combining the strengths of both methods. The technique can also adapt to changing encryption needs, improve the encryption efficiency, and resist attacks that traditional encryption techniques may fall prey to.	It may be computationally expensive and require specialized hardware and expertise to implement.

Consider two scenarios: the training and operating phase. In both cases, the network will most likely require a simple image dataset (see [31] for complete details). Typically, simple images are used to train the network. The original plain images should be decrypted to train the network, even if the image collection is encrypted. The individual who trains the network is referred to as a trainer in this context. The data holder who holds the training dataset is frequently not the same as the trainer. The data holder cannot then provide the dataset to the trainer using those two existing approaches, because doing so would violate the data holder's privacy policy. In the operational phase, the scenario is similar to that in the training phase. To detect or classify an object, the network requires a basic image. The images should be decrypted for the network even if the encrypted images are stored in the surveillance system. In this way, the operator who runs the networked surveillance system may always examine the original plain photographs. As a result, a new picture encryption challenge is presented here. The fundamental difference from the existing image encryption challenge is that the encrypted images have desired qualities. The algorithm should encrypt images against both humans and networks in the present image encryption challenge. In the image encryption challenge discussed here, encoded images should be encrypted for humans while the network can be trained on encoded images. This type of encryption is known as learnable image encryption. Learnable image encryption is capable of encrypting images for human use. It means that the data owner can give their dataset while being compliant with the privacy policy. Trainers can use encrypted photos to train directly. The development of networks is extremely beneficial because the data holder and the trainer can avoid privacy concerns. The learnable image encryption is also effective during the operation phase. Encrypted images are used to train the network. As a result, without decrypting original plain photos, the network can recognize or classify objects using directly encrypted images.

3. System Model

This section discusses a novel technique of face feature encryption with image optimization using cryptography and deep learning architectures. Here, the input face image has been processed and mapped using optical chaotic maps that are utilized for efficient encryption and decryption of the image. Then, the secure crypto general adversarial neural network was developed for encryption and the decryption method with image optimization. The overall proposed method is represented in Figure 2.

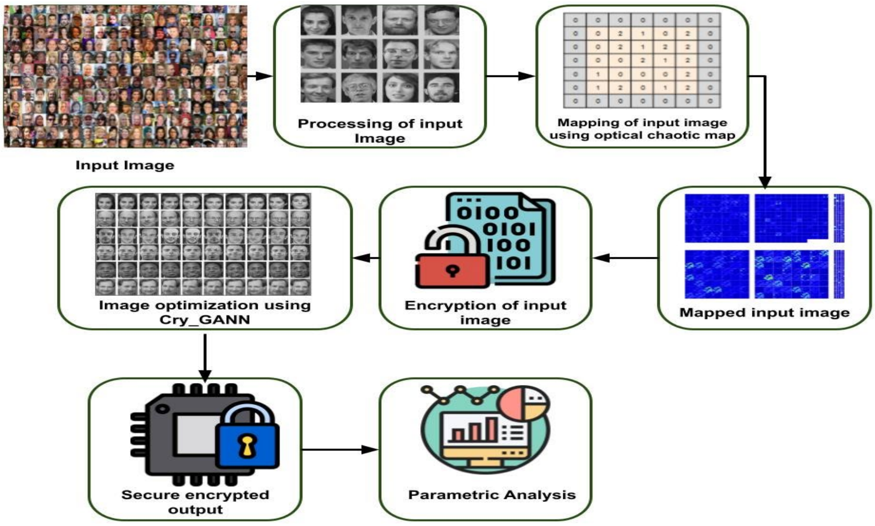

Input Image
Processing of input Image
Mapping of input image using optical chaotic map
Mapped input image
Encryption of input image
Image optimization using Cry_GANN
Secure encrypted output
Parametric Analysis

Figure 2. Overall proposed architecture.

3.1. Digital Optical Chaotic Mapping (Op-Ch_M)-Based Digital Image Encryption Technique

Xiong et al. offer a new chaotic map-based digital picture encryption technique. The surname initials of the author, abbreviated as XZQ, are used for brevity. The XZQ algorithm's encryption phases are listed below.

Phase 1. Select chaotic mapping to produce a chaotic sequence with beginning specifications of x_0, and a number sequence of = (n_1, n_2 , $\mathrm{n}_3 \ldots$, n_k), $0 < \mathrm{n}_\mathrm{i} < \mathrm{n}$ and $\sum_{i=1}^{k} n_i = n$ where n is the picture row size.

Phase 2. Create a double-precise chaotic sequence $\{x_1, x_2, \ldots, x_n\}$ using a chaotic mapping $f(x) = \mu x(1-x)$ arranging n items in the real sequence set $\{x_1, x_2, \ldots, x_n\}$ in increasing order to generate a systematic sequence $\left\{x'_1, \overline{x'_2, \ldots, x'_n}\right\}$ to produce a permutation address set $\{t_1, t_2, \ldots, t_n\}$; at this point, t_i numbers in $\{1, 2, \ldots, n\}$; commute pixels in first row based on $\{t_1, t_2, \ldots, t_n\}$, namely transposing pixels $t_i, i = 1, 2, \ldots, n$.

Phase 3. Set $x_1 = x_{n+n}$, and redo the process in Phase 2 in residual rows; $n_i = n_k$, repeatedly utilize δ from n_1 reintroduced δ.

Phase 4. Perform the same modifications in the image's rows to $L_G = \min_G \left(E_{x\sim pdata(x)} \log(1 - D(G(x))\right))$ to complete image encryption.

G network starts with a convolution stage to encode and compress pictures spatially, and useful characteristics retrieved in this phase are then utilized in the transformation that follows. Finally, a 7 × 7 convolution kernel exports the forecast. Furthermore, the decryption network F has a similar structure as encryption network G; encryption network G has successfully converted original patient images. Therefore, encrypted network G's loss LG is given by Equation (1):

$$\mathrm{L}_{\mathrm{reconstruction}} = \mathrm{E}_{\mathrm{x}\sim\mathrm{p}_{\mathrm{data}\ (\mathrm{x})}} \| \mathrm{Y} - \mathrm{X} \|_1 \tag{1}$$

G stands for an encryption network, and D stands for a discriminator network. G loss aims to reduce discriminator network D's success rate in detection ciphertext produced by encryption network G. Aside from encryption, another suggested technique is to make certain that the restored image retains the original image's texture data even when it is encrypted. Reconstruction loss estimates dissimilarity between G(x) and the original image for every image x from domain X, x → G(x) → F(G(x)) ≈ x. L is calculated using Equation (2):

$$\begin{aligned} \mathrm{L}_{\mathrm{reconstruction}} &= \mathrm{E}_{\mathrm{x}\sim\mathrm{p}_{\mathrm{data}\ (\mathrm{x})}} \| \mathrm{Y} - \mathrm{X} \|_1 = \mathrm{E}_{\mathrm{x}\sim\mathrm{p}_{\mathrm{data}\ (\mathrm{x})}} \sum_{i=1}^{\mathrm{n}} |\mathrm{y}_\mathrm{i} - \mathrm{x}_\mathrm{i}| \\ &= \mathrm{E}_{\mathrm{x}\sim\mathrm{p}_{\mathrm{data}\ (\mathrm{x})}} (|\mathrm{y}_1 - \mathrm{x}_1| + \ldots + |\mathrm{y}_\mathrm{i} - \mathrm{x}_\mathrm{i}|) \end{aligned} \tag{2}$$

The primary value of the 5D conservative chaotic method is evaluated by using pseudo-random sequence LCon based on the technique given below:

$$\begin{cases} \mathrm{x0} = \frac{\sum_{l=1}^{x} \mathrm{Lcon}(l)}{\omega_0} + \alpha_0 \\ \mathrm{y0} = \sum\limits_{l=1}^{n} \mathrm{Lcon}(l) \times \psi_0 + \beta_0 \\ \mathrm{u0} = \frac{\sum_{l=1}^{x} \mathrm{Lcon}(l) \times 0.48}{\varphi_0} + \mathrm{y}_0 \\ \mathrm{v0} = \frac{\log 2 \sum_{l=1}^{x} \mathrm{Lcon}(l) \times 0.48}{\varphi_0} + \delta_0 \end{cases}$$

where $n \in N$, and $n < d_0$. $I = 1, 2, \ldots, n$ is the index value. Initial control specifications are: $\alpha_0 = -0.16$, $\beta_0 = 5.52$, $\gamma_0 = -2.24$, $\delta_0 = -1.2$. $\varepsilon_0 = -0.3$. Initial scale coefficients are $\omega_0 = 40318$, $\psi_0 = 0.0004$, $\varsigma_0 = -2176$, $\phi_0 = -15140$, $\varphi_0 = 8.667$. By Equation (13), the initial values of the 5D conservative chaotic method are evaluated as $x_0 = 0.2536$, $y_0 = 7.0021$, $z_0 = -2.0216$, $u_0 = 2.5102$, $v_0 = -0.7109$.

Phase 2. Arbitrary sequences X, Y, Z, U, V are given by iterating the 5D conservative chaotic method, and arbitrary matrices RM_1, RM_2, K_1, K_2, K_3 are obtained by transformation:

$$\begin{cases} \mathrm{RM_1} = \bmod\left(\mathrm{X} \times 10^{15}, 256\right) \\ \mathrm{RM_2} = \bmod\left(\mathrm{Y} \times 10^{15}, 256\right) \\ \mathrm{K_1} = \bmod\left(\mathrm{Z} \times 10^{15}, 256\right) \\ \mathrm{K_2} = \bmod\left(\mathrm{U} \times 10^{15}, 256\right) \\ \mathrm{K_3} = \bmod\left(\mathrm{V} \times 10^{15}, 256\right) \end{cases}$$

RM_1 and RM_2 are arbitrary stage masks utilized for optical encryption channels, and K_1, K_2, K_3 are the keys to the digital diffusion channel's encryption.

Phase 3. Optical encryption is performed on the image with low-bit scrambling PL_{P2} to obtain low-bit encrypted image EL_{P2}. The encryption technique is described by the equation below:

$$\begin{gathered} \mathrm{EL_{P_2}(x,y)} = \\ \mathrm{F_{P_2} FP_2} \left\{ \mathrm{FP_n F^p p_y} \left[\mathrm{PL_{p_2}(x,y)} \times \mathrm{RM_1(u,v)} \right] \times \mathrm{RM_2(u,v)} \right\} \end{gathered}$$

where $F^{P_{x2}} F^{P_y}$ is fractional Fourier transform through order p_x for the x-axis and order p_y for the y-axis. The high-bit scrambled image with dynamic adaptive inverse diffusion PHP1 yields EHP1.

$$\begin{cases} \mathrm{EH_{P1}}(\tau) = \mathrm{bitxor}(\mathrm{K_1}(\tau), \mathrm{PH_{P1}} - \mu) \\ \mathrm{EH}(\tau + 1) = \\ \mathrm{bitxor\ (\ bitxor\ (EHP_{P1}}(\tau), \mathrm{PH_{P1}}(\tau + 1), \mathrm{K_2}(\tau + 1))) \\ \mathrm{EH_{P1}}(\tau - 1) = \\ \mathrm{bitxor\ (\ bitxor\ (EH} \\ \mathrm{P_1}(\tau), \mathrm{PH_{P1}}(\tau - 1), \mathrm{K_2}(\tau - 1))) \end{cases}$$

where μ is the dynamic diffusion control specification fixed by the user and τ is the dynamic diffusion direction control specification,

$$\tau = \bmod\left(\frac{\sum\limits_{1=1}^{M} \sum\limits_{j=1}^{N} \left(\frac{\mathrm{K_2} + 255}{3} \right)}{3 \times \mathrm{M} \times \mathrm{N}} \times 10^{16}, \mathrm{M} \times \mathrm{N} \right), \tau \in (1, \mathrm{M} \times \mathrm{N})$$

The choice to balance data in two ciphertext pictures can be made based on application needs. The option to delete the 4-bit information when the communication proportion is higher can also be considered. If the image's details are sought, it is essential to stabilize the data of an image with a high 4-bit ciphertext and an image with a low 4-bit ciphertext to obtain C1 and C2. In this example, the image balance approach is illustrated.

$$\mathrm{C}_1,\mathrm{C}_2 \leftarrow \begin{cases} \xi 3 = \mathrm{EH}_{\mathrm{P1}}(\mathrm{x}_1(\mathrm{i}),\mathrm{y}_1(\mathrm{j})) \\ \mathrm{EH}_{\mathrm{P1}}(\mathrm{x}_1(\mathrm{i}),\mathrm{y}_1(\mathrm{j}))= \mathrm{EL}_{\mathrm{p2}}(\mathrm{z}_1(\mathrm{i}),\mathrm{w}_1(\mathrm{j})) \\ \mathrm{EL}_{\mathrm{p2}}(\mathrm{z}_1(\mathrm{i}),\mathrm{w}_1(\mathrm{j}))= \xi 3 \end{cases}$$

Discriminator network D seeks to distinguish between translated samples by maximizing discriminator network D's classification accuracy, which is the inverse of the encryption network G's goal Equation (3):

$$\mathrm{L_D}= \mathrm{E}_{\mathrm{x}\sim\mathrm{pdata(x)}}\mathrm{logD(x)} + \mathrm{E}_{\mathrm{x}\sim\mathrm{pdata(x)}}\mathrm{log}(1-\mathrm{D}(\mathrm{G}(\mathrm{x}))) \tag{3}$$

The private key for encryption is the final specifications of network *G*. In contrast, the private key for decryption is the final parameters of network *F*. The following is the procedure for producing a privacy key: For encryption, every convolutional layer's parameters are initially arbitrarily initialized as given: $W_n = \mathrm{random}\left[w_{n,1}, w_{n,2}, \ldots, w_{n,j}, \ldots\right]$ where w_n is the nth convolutional layer. As a result, the encryption privacy key W is made up of all specifications of every convolutional layer, which is described as follows: $W = \mathrm{consist}$ $[W_1, W_2, \ldots, W_n, \ldots]$.

In addition to forward propagation, the BP method transfers network loss between convolutional layers. Improve performance by updating the parameters in each layer. The gradient descent is defined by Equation (4):

$$\begin{aligned}
\theta_j &= \theta_j - \alpha \vee \mathrm{J}(\theta) \\
&= \theta_j - \alpha \frac{\delta}{\theta_j}\mathrm{J}(\theta) \\
&= \theta_j - \alpha \frac{\delta}{\theta_j}\frac{1}{2\,\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}\left(\mathrm{h}_\theta\left(\mathrm{x}^{\mathrm{i}}\right)-\mathrm{y}^{\mathrm{i}}\right)^2 \\
&= \theta_j - \alpha \frac{1}{2\,\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}\frac{\delta}{\theta_j}\left(\mathrm{h}_\theta\left(x^{\mathrm{i}}\right)-\mathrm{y}^{\mathrm{i}}\right)^2 \\
&= \theta_j - \alpha \frac{1}{2\,\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}2\frac{\delta}{\theta_j}\left(\mathrm{h}_\theta\left(\mathrm{x}^{\mathrm{i}}\right)-\mathrm{y}^{\mathrm{i}}\right)\left(\frac{\delta}{\theta_j}\left(\mathrm{h}_\theta\left(\mathrm{x}^{\mathrm{i}}\right)-\mathrm{y}^{\mathrm{i}}\right)\right) \\
&= \theta_j - \alpha \frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}\left(\mathrm{h}_\theta\left(\mathrm{x}^{\mathrm{i}}\right)-\mathrm{y}^{\mathrm{i}}\right) \times \left(\sum_{\mathrm{i}=1}^{\mathrm{n}}\frac{\delta}{\theta_{\mathrm{i}}}\theta_{\mathrm{i}}\mathrm{x}_{\mathrm{i}} - \frac{\delta}{\theta_{\mathrm{i}}}\mathrm{y}^{\mathrm{i}}\right)
\end{aligned} \tag{4}$$

The procedure of creating a privacy key for decryption is the same as producing a privacy key for encryption, excluding that the decryption network's initial input becomes the encryption network's projected output. Furthermore, the reconstruction loss is the loss of the decryption network, as shown in Equation (5).

$$\mathrm{L}_{\mathrm{reconstruction}} = \mathrm{E}_{\mathrm{x}\sim\mathrm{p}_{\mathrm{data}}\,(\mathrm{x})}\sum_{\mathrm{i}=1}^{\mathrm{n}}|\mathrm{F}(\mathrm{P}(\mathrm{x}_{\mathrm{i}}))-\mathrm{O}(\mathrm{x}_{\mathrm{i}})| \tag{5}$$

The encryption algorithm is as follows:

1. Calculate the *H* value by extracting the characteristic value of the image to be encrypted.

$$x_i = mod\Big(\Big(abs(x_i) - floor(abs((x_i)))\Big) \times 10^{44}, 256\Big) i = 1,2,3,4$$

2. To carry out the process, utilize initial chaotic value x_0 and H value, producing initial value x 0′ utilized in scrambling chaotic sequence $\{C_i, i = 0, 1, \ldots, M * N - 1\}$ as explained in Figure 3.

3. Arrange the chaotic sequence C_i in descending order; the resulting sequence is $C_i^{'}$. Calculate mapping matrix A for converting C_i to $C_i^{'}$, for example, $C_i^{'} = A * C_i$.
4. To obtain the final encrypted image $G_0^{'}$, utilize matrix A to scramble the image G_0 according to the pixel location. $G_0^{'}.G_0^{'} = A * G_0$.
5. Decryption method: Extract the characteristic value of the image to be decoded.

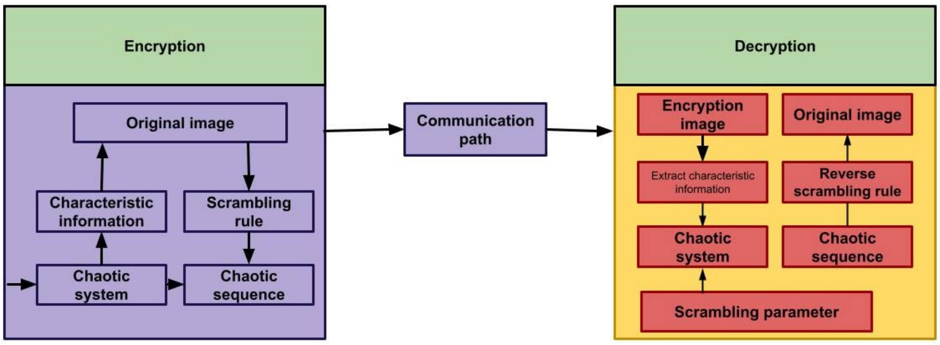


Figure 3. Schematic diagram for the chaos-based scrambling algorithm.

Step 1. Calculate the hyperchaotic system's generated random sequence using Equations (6) and (7).

$$\mathrm{x_i}= \mathrm{mod}\Big(\Big(\mathrm{abs(x_i)} - \mathrm{floor(abs((x_i)))} \times 10^{44}, 256\Big)\mathrm{i} = 1, 2, 3, 4 \tag{6}$$

$$\text{Obviously, } x_i \in [0, 255]$$

$$\overleftarrow{\mathrm{x}}_1= \mathrm{mod}((\mathrm{x_1+x_2+x_3+x_4}), 4) \tag{7}$$

Step 2. Encrypt the acquired row–column permutation matrix and select the appropriate combination from Table 2 based on $x_1[0.3]$.

$$\begin{gathered}\mathrm{C_{3\times(i-1)+1} = P^{rc}{}_{3\times(i-1)+1} \oplus D_{x_1}} \\ \mathrm{C_{3\times(i-1)+2} = P^{rc}{}_{3\times(i-1)+1} \oplus D_{x_2}}, \\ \mathrm{C_{3\times(i-1)+3} = P^{rc}{}_{3\times(i-1)+1} \oplus D_{x\,_3}}, \end{gathered} \tag{8}$$

where D_{x_1},D_{x_2}, and D_{x_3} are given in Equation (9):

$$\begin{gathered}\mathrm{D_{x_1}}= \mathrm{mod}\Big(\Big(\mathrm{B_{x_1} \oplus C_{3\times(i-1)+1}}, 256\Big), \\ \mathrm{D_{x_2}}= \mathrm{mod}\Big(\Big(\mathrm{B_{x_1} \oplus C_{3\times(i-1)+2}}, 256\Big), \\ \mathrm{D_{x_3}}= \mathrm{mod}\Big(\Big(\mathrm{B_{x_1} \oplus C_{3\times(i-1)+3}}, 256\Big), \end{gathered} \tag{9}$$

Table 2. Comparative analysis between the proposed and existing techniques.

Datasets	Techniques	PSNR	SSIM	RMSE	MAP	Encryption Speed
ImageNet dataset	CNN	88	78	79	61	85
	IEA	90	82	71	55	86
	Cry_GANN_OChaMap	92	85	68	52	88
LFW	CNN	85	83	69	65	81
	IEA	88	85	65	61	82
	Cry_GANN_OChaMap	90	89	61	59	89

Obviously, $D_x \in [0, 255]$, where t = 1, 2, ... relates i-th hyperchaotic iteration; signifies XOR, $\mathrm{P}_i, i = 1, 2, \ldots, M \times N$ relates scrambled image's pixel value; B_{x_1}, B_{x_2}, and B_{x_1} reflect the corresponding combinations in Table 2 selected based on $\overline{x_1}, \mathrm{C}_i, i = 1, 2, \ldots, M \times N$. Step 3. The encryption procedure is complete if all plaintexts have been encrypted; otherwise, proceed to Step 1. The encryption and decryption processes are comparable. First, build the same hyperchaotic sequence with the same parameters and beginning values, but replace it with Equation (10), as follows:

$$\begin{aligned} \mathrm{P}^{\mathrm{rc}}{}_{3\times(\mathrm{i}-1)+1} &= \mathrm{C}_{3\times(\mathrm{i}-1)+2} \oplus \mathrm{D}_{\mathrm{x}_2}, \\ \mathrm{P}^{\mathrm{rc}}{}_{3\times(\mathrm{i}-1)+1} &= \mathrm{C}_{3\times(\mathrm{i}-1)+3} \oplus \mathrm{D}_{\mathrm{x}_3}, \end{aligned} \tag{10}$$

Then, according to $\{r_i, i = 0, 1, \ldots, M-1\}$ and $\{c_j, j = 0, 1, \ldots, N-1\}$, the matrix is inversely transformed, and the original image is restored as shown in Figure 4.

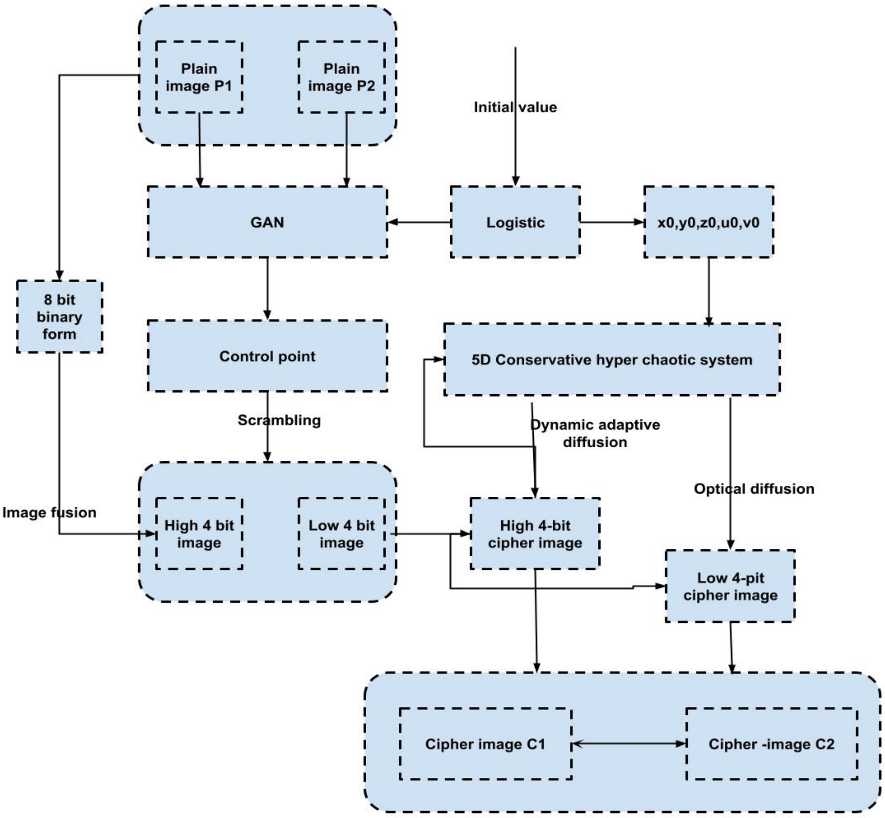


Figure 4. Flowchart of the encryption algorithm.

In contrast to the typical method of training, i.e., two methods as a generator and a discriminator, three NNs are used here. A pair of NNs act as generators, while a third acts as a modified discriminator. Three NNs will be:

6. Encryptor: Plaintext and a shared key, both in binary sequence, are used to produce encrypted text.
7. Decryptor: The encrypted text is used as input, and the shared key are used to produce an output of decrypted text.
8. Eavesdropper: This only accepts the encrypted text as input, which means it intercepts text and decrypts it without the shared key.

The general adversarial network (GAN) has the following layers; the architecture of all three NNs is the same as the subsequent layers:

- Dense layer that is fully linked;
- Flatten layer;
- Convolutional layer.

We employed one dense, four convolutional, and one flattened layer. Strided convolution is utilized to replace the pooling layer. We employed strided convolutions instead of immediate downsampling. Activation functions used: binary sequences, 0 and 1, are used for encryption. To standardize the output of each layer in $[1,\ 1]$, the tanh initiation is utilized, but the last layer is utilized for sigmoid activation.

The output of A on inputs FF and Key is represented by $\mathrm{A}(\omega_{\mathrm{A}},\ \mathrm{FF},\ \mathrm{Key})$, the output of the Server is represented by $\mathrm{S}(\omega_{\mathrm{Server}}\ ,\ \mathrm{EFF},\ \mathrm{Key})$, and the output of B on input C is represented by $\mathrm{B}(\omega_{\mathrm{B}},\ \mathrm{EFF})$.

The distance function d is also incorporated into the facial feature at the same time. This study uses L_2 distance $d_2\left(\mathrm{FF},\mathrm{FF}^{'}\right) = \sqrt{\sum_{i=1,N}(\mathrm{FF}_1^n - \mathrm{FF}_2^n)^2}$ for a specific operation, where N is the length of the facial feature. The loss function for each instance of B is defined by Equation (11):

$$\mathrm{L_B}(\omega_{\mathrm{A}},\ \omega_{\mathrm{B}},\ \mathrm{FF},\ \mathrm{Key}) = \mathrm{d}(\mathrm{FF},\ \mathrm{B}(\omega_{\mathrm{B}},\ \mathrm{A}(\omega_{\mathrm{A}},\ \mathrm{FF},\ \mathrm{Key}\ ))) \tag{11}$$

$\mathrm{L_B}(\omega_{\mathrm{A}},\omega_{\mathrm{B}},\ \mathrm{FF},\ \mathrm{Key})$ reflects the inaccuracy of B when facial features are FF and key are in Equation (12).

$$\mathrm{L_B}(\omega_{\mathrm{A}},\omega_{\mathrm{B}}) = \mathrm{E_{FF,\ Key}}(\mathrm{d}(\mathrm{FF},\ \mathrm{B}(\omega_{\mathrm{B}},\ \mathrm{A}(\omega_{\mathrm{A}},\ \mathrm{FF},\ \mathrm{Key})))) \tag{12}$$

This research acquires the "best B" by reducing that loss, as shown in Equation (13):

$$\mathrm{O_B}(\omega_{\mathrm{A}},\ \omega_{\mathrm{Server}}\ ) = \mathrm{argmin}_{\omega_{\mathrm{B}}}(\mathrm{L_B}(\omega_{\mathrm{A}},\ \omega_{\mathrm{B}})) \tag{13}$$

In a similar vein, this paper constructs a sample Server reconstruction error and applies it to a distribution of face characteristics and keys by Equation (14):

$$\begin{gathered}\mathrm{L_{server}}\ (\omega_{\mathrm{A}},\ \omega_{\mathrm{Server}}\ ,\ \mathrm{FF},\ \mathrm{Key}\ ) = \mathrm{d}(\mathrm{FF},\ \mathrm{Server}(\omega_{\mathrm{Server}}\ ,\ \mathrm{A}(\omega_{\mathrm{A}},\ \mathrm{FF},\ \mathrm{Key}\ ),\ \mathrm{Key}\ )) \\ \mathrm{L_{Server}}\ (\omega_{\mathrm{A}},\ \omega_{\mathrm{Server}}\ ) = \mathrm{E_{FF,\ Key}}\ (\mathrm{d}(\mathrm{FF},\ \mathrm{B}(\omega_{\mathrm{Server}}\ ,\ \mathrm{A}(\omega_{\mathrm{A}},\ \mathrm{FF},\ \mathrm{Key}))))\end{gathered} \tag{14}$$

The optimal values of L_{server} and $\mathrm{L_B}$ are combined in this study to define the Server and loss function of A by Equation (15):

$$\mathrm{L_{AServer}}(\omega_{\mathrm{A}},\ \omega_{\mathrm{Server}}\ ) = \mathrm{L_{Server}}\ (\omega_{\mathrm{A}},\ \omega_{\mathrm{Server}}\ ) - \mathrm{L_B}(\omega_{\mathrm{A}},\ \mathrm{O_B}(\omega_{\mathrm{A}})) \tag{15}$$

This combination illustrates the aim of A and Server to reduce Server rebuild faults while increasing "optimal B" rebuild errors. However, the following research discusses beneficial alternatives.

By reducing $\mathrm{L}_{\mathrm{A\ Server}}\ (\omega_{\mathrm{A}}, \omega_{\mathrm{Server}}\ )$, this study obtains "Best A and Server" by Equation (16):

$$(\mathrm{O}_{\mathrm{A}},\ \mathrm{O}_{\mathrm{Server}}\ )= \mathrm{argmin}_{(\omega_{\mathrm{A}}\omega_{\mathrm{Server}}\ )}(\mathrm{L}_{\mathrm{AServer}}(\omega_{\mathrm{A}},\ \omega_{\mathrm{Server}}\ )) \tag{16}$$

3.2. Secure Crypto General Adversarial Neural Network

In this subsection, we provide the following algorithm for securing the crypto general adversarial neural network.

```
Require:
$c$, clipping parameter, $m_t$, batch size. $\beta_1\beta_2$, hyperparameters parameters; $n_{\text{critic}}$ amount of generator iterations per critic iteration.
while $\theta$ has not joined do
for $t = 0, \ldots, n_{\text{cricic}}$ do 3; for $i = 1,\ 2,\ 3, \ldots, m$ do
Sample $\left\{x^{(i)}\right\}_{i=1}^{m} - P_r$.
Sample $\left\{2^{(i)}\right\}_{i=1}^{in} - p(z)$.
a random number $\varepsilon \sim U[0,1]$.
 $\overleftarrow{x} \leftarrow G_\theta(z)$
 $\grave{t} \leftarrow \varepsilon x + (1-c)\overleftarrow{x}$
 $L^{(i)} \leftarrow D_w\left(\overrightarrow{x}\right) - D_w(x) + \lambda\left(\|\nabla_{\grave{x}} D_w(\grave{P})\|_2 - 1\right)^2$
 $g_w \leftarrow \nabla_w\left[\frac{1}{w}\sum_{i=1}^{m} f_w\left(x^{(i)}\right) - \frac{1}{m}\sum_{i=1}^{w} f_w\left(g_e\left(z^{(i)}\right)\right)\right]$
   $w \leftarrow w + \alpha \cdot \mathrm{RMSProp}(w, 8w)$
   $w \leftarrow \mathrm{clip}(w, -c, c)$
  end for
  end for
   $w \leftarrow \mathrm{Adam}\left(\nabla_w \frac{1}{w}\sum_{i=1}^{m} L^{(i)}, w, \alpha, \beta_1, \beta_2\right)$
  end for
  Sample $\left\{z^{(i)}\right\}_{i=1}^{m} \sim p(z)$.
   $g_\theta \leftarrow -\nabla_\theta \frac{1}{m}\sum_{i=1}^{m^1} f_w\left(g_\theta\left(z^{(i)}\right)\right)$
   $\theta' \leftarrow \theta - \alpha \cdot \mathrm{RMSProp}(\theta, g_\theta)$
   $\theta \leftarrow \mathrm{Adam}\left(\nabla_\theta \frac{1}{m}\sum_{i=1}^{\mathrm{m}} -D_\omega(G_\theta(z)), \theta'; \alpha, \beta_1, \beta_2\right)$
  end while
```

A GAN consists of two neural networks: a generator network and a discriminator network. The generator network is trained to generate new samples that are similar to a target distribution, while the discriminator network is trained to distinguish between the generated samples and real samples from the target distribution.

In this algorithm, the generator network is represented by G_θ, and the discriminator network is represented by D_w. The generator network is trained to generate samples, denoted by x, from a noise distribution, $p(z)$, while the discriminator network is trained to classify samples as either real (from the target distribution P_r) or fake (generated by the generator network).

The training process alternates between updates to the generator network and the discriminator network. During each generator update, the generator network generates a batch of samples from the noise distribution, and the discriminator network is updated to classify these samples as fake. During each discriminator update, the discriminator network is updated to classify a batch of real samples and a batch of fake samples generated by the generator network. The algorithm also uses various hyperparameters, such as the clipping parameter c, the batch size m, and the Adam optimization parameters α, β_1, and β_2. These hyperparameters set and tuned through experimentation to achieve the best performance for the specific task and dataset. The root mean squared propagation (RMSProp) optimization algorithm is also used to update the weights of the networks. The training process continues until the generator network has converged. The threshold value in the algorithm is the "clipping parameter" (c). It appears in the line "$w \leftarrow clip(w, -c, c)$", which sets the values of w to be within the range $(-c,\ c)$. The value of c is specified as an input to the algorithm and determines the maximum magnitude for the values of w.

A GAN's purpose is to predict the possible distribution of existing data and produce new data samples with the same distribution. Generator G's ability to produce samples is improved by creating a minimax confrontation procedure between it and discriminator D. The GAN's main purpose is to create a generator G from real-world data X. The model's objective function is given by Equation (17):

$$\min_{\mathrm{G}} \max_{\mathrm{D}} = \mathrm{E}_{\mathrm{x}\sim\mathrm{P}_\mathrm{d}}[(\log(\mathrm{D}(\mathrm{x})] + \mathrm{E}_{\mathrm{z}\sim\mathrm{P}_\mathrm{z}}[(1-\mathrm{D}(\mathrm{G}(\mathrm{z})))] \quad (17)$$

The distribution and the appearance of an image become increasingly similar as those of the cover image when a discriminator is added to the steganography system. For example, the following is the loss function given by Equation (18):

$$\mathrm{L}_{\mathrm{disc}} = \mathrm{E}_{\mathrm{c}\sim\mathrm{P}_\mathrm{c}}[\log\ \mathrm{D}(\mathrm{c})] + \mathrm{E}_{\mathrm{c}\sim\mathrm{P}_\mathrm{c},\mathrm{s}\sim\mathrm{P}_\mathrm{s}}[\log(1-\mathrm{D}(\mathrm{H}(\mathrm{c},\mathrm{s}))] \quad (18)$$

Here, P_c and P_s are covers and secret image distributions, and $H\ (c,\ s)$ is the steg image produced by the generator. Throughout the training operation, the hiding network and the extracting network are optimized to minimize the original secret image s and the loss of image-extracted secret image s 0. To do this, we propose a novel cost function that is reduced to enhance the method, which is given by Equation (19):

$$\underset{\mathrm{H}_\theta}{\mathrm{argmin}} \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} (1 - \mathrm{SSIM}(\mathrm{c}_\mathrm{i}, \mathrm{H}_\theta(\mathrm{c}_\mathrm{i}, \mathrm{s}_\mathrm{i})) \quad (19)$$

3.3. Security Analysis

Both the encryption and decryption networks include 24 levels, with a total of 2,757,936 parameters for each network. A deeper resnet-50 design is used for the ROI mining network. The ROI-mining network's structure is shown in Table 2. Chest X-rays [45] are the dataset. The proposed solution is implemented on Nvidia Giga Texel Shader eXtreme (GTX) 2080Ti graphics card. Each epoch of the model takes about 10 minutes to train the network.

3.3.1. Analysis of Key Space

The difficulty of an exhaustive attack is computed by the size of the key space. The number of specifications for the DL network is a key space of the proposed encryption method in this study, with an overall of 2,757,936 specifications in experimentations. Every specification or key is a 32-bit floating point value amid 0 and 1, which are written as a decimal integer with 10 significant digits in the computer. As a result, the encryption model's key space can be stated as (1010) 2757936. Attackers will find it difficult to break down the system, and it will be able to efficiently resist attacks.

3.3.2. Key Randomness Analysis

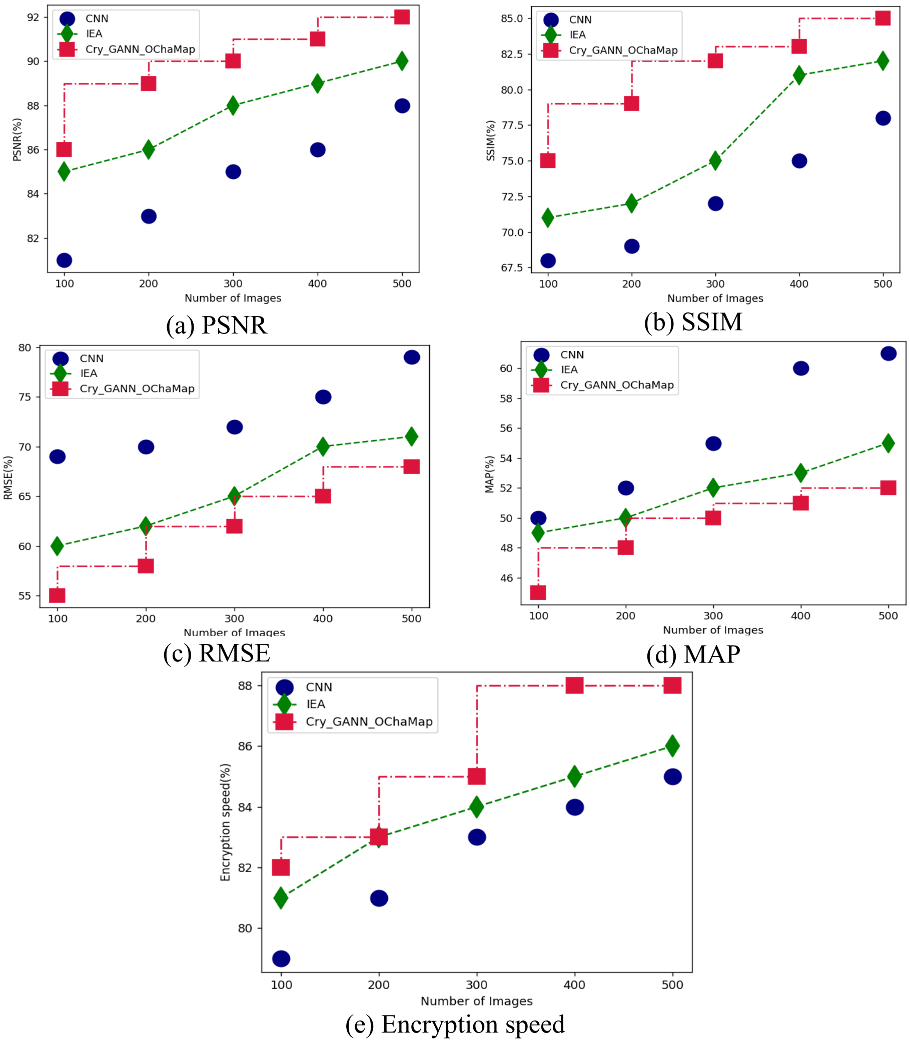
With the same conditions, the encryption network is trained four times. As a result, the parameters of these four networks, namely Key A, B, C, and D, are used as encryption keys. These four photos are unmistakably distinct. The SSIM index between various photos is usually less than 0.1, indicating that there is very little similarity between them. According to the experiment, the privacy keys for the medical picture encryption network are completely distinct after each training since the neural network's parameters are randomly initialized. As a result of these differences, separate encrypted images are produced, each of which is processed using a different encryption network. The premise is that DL network training is inherently unstable. In different training, different initialization parameters might lead to different end parameters. It is shown that the proposed technique is similar to OTP and that it may be classified as an OTP technique.

3.3.3. Key Sensitivity Analysis

DL models, unlike typical encryption systems, spread errors among layers. A 3 × 3 convolution kernel is used to send the lth pixel in the Nth layer feature map to a nearby pixel in (N + 1)th layer during the convolution process. When a feature point is incorrect, it is transferred to the next layer's 3 × 3 feature points. The inaccuracy of feature points will grow by two pixels for every layer as the depth of the convolutional network increases. This inaccuracy grows exponentially with the superposition of the deconvolution process in the upsampling process. The attacker is assumed to have the most privacy keys in this experiment. Only around 5% of crucial specifications are changed, which is considered an unknown part. The encrypted image is then sent to the network with new specifications, and the network is unable to convert the ciphertext image back to the original. This means that even if only 5% of specifications are modified, the privacy key will fail to properly encrypt or decode the medical image. In other words, breaking the proposed technique requires attackers to estimate at least 95% of correct key specifications in a key space containing (1010) 2757936.

3.3.4. Histogram Analysis

To calculate the performance of the suggested encryption network, the original image and the encrypted image are given in Figure 5c. The pixel distribution in the original image and the encrypted image is considerably varied, according to the experiment. The original chest X-ray image's pixel histogram has 57,600 * (240 * 240) pixels overall, with more than 30,000 pixels having a value of 0, and greater than 5000 pixels having a value of 255. The original image's pixel dissemination is fairly intense. The distribution of encrypted medical photos, on the other hand, is more uniform, which aids statistical analysis.



(a) PSNR

(b) SSIM

(c) RMSE

(d) MAP

(e) Encryption speed

Figure 5. Comparative analysis between the proposed and existing methods for ImageNet dataset in terms of (**a**) PSNR, (**b**) SSIM, (**c**) RMSE, (**d**) MAP, and (**e**) Encryption speed.

3.3.5. Entropy Analysis

In contrast to statistical attacks, encrypted image data entropy is considered a significant measurable dimension for methods to defend. Image data entropy is a statistical property of an image's grayscale distribution. The encrypted image should resemble random noise in appearance, the grayscale distribution should be uniform, and the anticipated value should equal 8.

3.3.6. Security Analysis under Various Adversary Models

Tests are being carried out to determine if an attacker can produce a key under three various adversary methods. Hidden Factors Leakage: Four potential network structures are investigated in this experiment: network A, network B, network C, and network D. The training environment remains unchanged. Utilizing trained network A, the original image is encrypted. To restore the original image, the ciphertext image is decrypted by a decryption network retrieved from network A, network B, network C, and network D.

3.3.7. Network Architecture Leakage

Different hidden factors are used in this experiment to train encryption networks with the same network configuration. All training circumstances remain constant.

3.3.8. Both Network Architecture Leakage and Hidden Factors

To produce networks A, B, C, and D, the network is trained four times in this experiment under the same hidden elements and training circumstances. Investigation compares the decryption performance of these four networks on identical ciphertext images to determine if specifications produced by each are unique. Using decryption keys B, C, and D, the image's gray value distribution differs significantly from that of the image decrypted with the help of decryption key A. It is seen that one network's encryption of a medical image prevents its decryption by utilizing specifications of another network under same training conditions. Even if method specifications are learned using the same network method and hidden factors, they will not be able to decode the image. Experiments reveal that even though both network design and hidden elements are leaked, and the network is trained under identical circumstances, the specifications of every network are completely dissimilar, i.e., secure keys.

4. Experimental Analysis

The experiments were carried out on a personal computer (PC) with an NVIDIA GeForce Tesla V100 32G graphics processing unit (GPU), Pytorch 1.1, and Python 3.7 as the experimental environment. The input dataset's image size is 32 × 32 pixels. The proposed method block size is set to 4. After block adaptation networks [7,8], pyramidal residual networks were created. Plain pictures, integrated cat map image encryption [9], a naive blockwise pixel shuffle, and the suggested technique are all compared here.

ImageNet dataset was used to gather 80,000 training photos and 10,000 test images for training network methods in this study. The Adam optimization technique was utilized to automatically modify the learning rate in the training phase to optimize method specifications. Hyper specifications were adjusted to 0.65 and 0.85, with the initial learning proportion set at 0.0001. The maximum number of training iterations was set to 250, while number of images per batch was set at 64. The Labeled Faces in the Wild (LFW) dataset was used for training and testing. This is a regularly used facial recognition test set. Because the face photographs are all taken from real-life situations, recognition difficulty is heightened, particularly due to elements such as different positions, expressions, age, lighting, and occlusion. Photos of the same individual are rather different. Multiple faces may appear in certain photographs. Only the center coordinate face is chosen as the goal in these multiface images, with the others being background interference. The LFW dataset contains 13,233 face pictures. Each photograph contains the name of the individual depicted. There are 5749 persons, and most of them have just one photo. Each photograph is 250 × 250 pixels in size; most are color images, but some are black and white portraits.

The comparative analysis between the proposed and existing techniques is conducted on the two datasets, ImageNet dataset and LFW, in terms of PSNR, SSIM, RMSE, MAP, and encryption speed. These matrices are defined below.

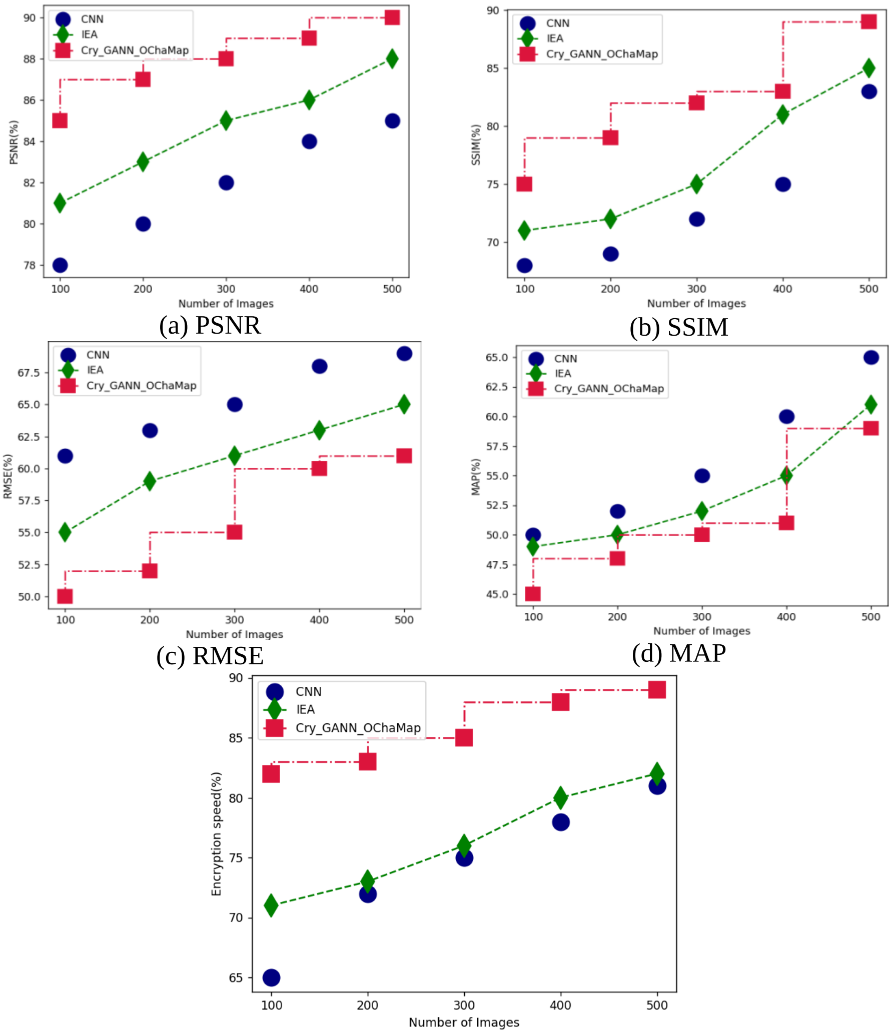
Peak signal-to-noise ratio, or PSNR, measures how well the original data and the reconstructed data match in terms of quality. It is frequently used to assess the effectiveness of data compression and reconstruction methods in the field of image and video processing.

A measure that evaluates the similarity between two pictures is called SSIM (structural similarity index). It is predicated on the idea that structural information, such as the association between neighboring pixels, is easily perceptible by humans due to their highly developed visual system. Images are compared for structural similarities using SSIM.

The discrepancy between two sets of data is measured by RMSE (root mean squared error). It is employed to evaluate the discrepancy between values that a model predicts and values that are observed. Information retrieval and classification algorithm performance are assessed using the MAP (mean average precision) measure. In the field of computer vision, it is frequently used to assess how well object detection methods are working.

PSNR (peak signal-to-noise ratio) is usually expressed in decibels (dB). SSIM (structural similarity index) is a dimensionless value that is typically expressed as a decimal value between -1 and 1, where a value of 1 indicates that the two images are identical. RMSE (root mean squared error) is typically expressed in terms of the pixel values of the images. MAP (mean average precision) is a dimensionless value that is usually expressed as a decimal value between 0 and 1. Encryption speed is typically expressed in terms of the number of operations per second that can be performed by the encryption algorithm.

Table 2 presents a comparative evaluation between suggested and existing methods in face image encryption based on DL architectures. Here, the ImageNet and LFW face datasets are compared when the proposed Cry_GANN_OChaMap and existing CNN and IEA methods are applied. The parametric analysis was carried out regarding PSNR, RMSE, SSIM, MAP, and encryption speed. Parametric analysis for the ImageNet and LFW datasets are displayed in Figures 5 and 6, respectively. Initially, for the ImageNet dataset, the proposed Cry_GANN_OChaMap obtained PSNR of 92%, RMSE of 85%, SSIM of 68%, MAP of 52%, and encryption speed of 88%, as shown in Figure 5a through (e) respectively; while the LFW dataset attained PSNR of 90%, RMSE of 89%, SSIM of 61%, MAP of 59% and encryption speed of 89%, as shown in Figure 6a through (e), respectively. From this analysis, it can be observed that the proposed technique attained optimal results in face encryption based on DL techniques.



(a) PSNR

(b) SSIM

(c) RMSE

(d) MAP

(e) Encryption speed

Figure 6. Comparative analysis between the proposed and existing techniques for the LFW dataset in terms of (**a**) PSNR, (**b**) SSIM, (**c**) RMSE, (**d**) MAP, and (**e**) Encryption speed.

The digital optic chaotic mapping technique is utilized as a preprocessing step for the proposed face feature encryption technique. Chaotic mapping is being used to process and map the input face image, which is then fed into the Crypto General Adversarial neural network for encryption and decryption. The use of chaotic mapping allows for efficient encryption and decryption of the image, and the combination of chaotic mapping with deep learning techniques allows for the development of a secure encryption and decryption method. The results demonstrate that the proposed technique for face encryption outperformed existing techniques.

5. Conclusions

This research proposed a novel technique in secure face image encryption based on DL architectures. Here, the input face image is processed and mapped using optical chaotic maps, which are utilized for efficient encryption and decryption of the image. Then, the secure Crypto General Adversarial neural network was developed for the encryption and decryption process with image optimization. The key is crucial to successful encryption and decoding in cryptography. The key's security determines information security. The standard image encryption scheme is vulnerable to key sharing and repudiation attacks. If the key is excessively long, it is difficult to remember and simple to lose. As times demanded, biometric encryption technology arose to address the issue of key security. The key is produced using an individual's biometrics, and then used with appropriate picture encryption methods to attain data encryption. Uniqueness, stability, nonaggression, and other criteria of biological features that can be encrypted should be met. Experimental analysis was carried out for various datasets; the proposed Cry_GANN_OChaMap technique resulted in PSNR of 92%, RMSE of 85%, SSIM of 68%, MAP of 52%, and encryption speed of 88%.

Future research will focus on how the system learns reliable encryption algorithms for asymmetric encryption. Additionally, it is thought that the encryption technique developed through adversarial training may be used for a larger range of data types, including audio and image data, in addition to character data security.

Author Contributions: M.A.: methodology, F.A.: Conceptualization, H.A.: formal analysis, E.A.: investigation. All authors have read and agreed to the published version of the manuscript.

Funding: The Taif University Deanship of Scientific Research provided funding for this endeavor. The authors want to express their gratitude to Taif University for providing funding for this study (project number: 20228).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors certify that they have no competing interests in relation to this research.

References

1. Ding, Y.; Wu, G.; Chen, D.; Zhang, N.; Gong, L.; Cao, M.; Qin, Z. DeepEDN: A Deep-Learning-Based Image Encryption and Decryption Network for Internet of Medical Things. *IEEE Internet Things J.* **2020**, *8*, 1504–1518. [CrossRef]
2. Wu, C.; Ju, B.; Wu, Y.; Xiong, N.N.; Zhang, S. WGAN-E: A Generative Adversarial Networks for Facial Feature Security. *Electronics* **2020**, *9*, 486. [CrossRef]
3. Li, X.; Jiang, Y.; Chen, M.; Li, F. Research on iris image encryption based on deep learning. *EURASIP J. Image Video Process.* **2018**, *2018*, 126. [CrossRef]
4. Man, Z.; Li, J.; Di, X.; Sheng, Y.; Liu, Z. Double image encryption algorithm based on neural network and chaos. *Chaos Solitons Fractals* **2021**, *152*, 111318. [CrossRef]
5. Maniyath, S.R.; Thanikaiselvan, V. An efficient image encryption using deep neural network and chaotic map. *Microprocess. Microsyst.* **2020**, *77*, 103134. [CrossRef]
6. Deng, Z.; Zhong, S. A digital image encryption algorithm based on chaotic mapping. *J. Algorithms Comput. Technol.* **2019**, *13*, 1748302619853470. [CrossRef]
7. Li, Q.; Wang, X.; Wang, X.; Ma, B.; Wang, C.; Xian, Y.; Shi, Y. A Novel Grayscale Image Steganography Scheme Based on Chaos Encryption and Generative Adversarial Networks. *IEEE Access* **2020**, *8*, 168166–168176. [CrossRef]
8. Chai, X.; Tian, Y.; Gan, Z.; Lu, Y.; Wu, X.-J.; Long, G. A robust compressed sensing image encryption algorithm based on GAN and CNN. *J. Mod. Opt.* **2022**, *69*, 103–120. [CrossRef]
9. Ding, Y.; Tan, F.; Qin, Z.; Cao, M.; Choo, K.-K.R.; Qin, Z. DeepKeyGen: A Deep Learning-Based Stream Cipher Generator for Medical Image Encryption and Decryption. *IEEE Trans. Neural Networks Learn. Syst.* **2021**, *33*, 4915–4929. [CrossRef]

10. Wu, C.; Chang, J.; Xu, X.; Quan, C.; Zhang, X.; Zhang, Y. Cryptoanalysis of the modified diffractive-imaging-based image encryption by deep learning attack. *J. Mod. Opt.* **2020**, *67*, 1398–1409. [CrossRef]
11. Zhou, L.; Xiao, Y.; Chen, W. Vulnerability to machine learning attacks of optical encryption based on diffractive imaging. *Opt. Lasers Eng.* **2020**, *125*, 105858. [CrossRef]
12. Feng, F.; Hu, J.; Guo, Z.; Gan, J.-A.; Chen, P.-F.; Chen, G.; Min, C.; Yuan, X.; Somekh, M. Deep Learning-Enabled Orbital Angular Momentum-Based Information Encryption Transmission. *ACS Photon.* **2022**, *9*, 820–829. [CrossRef]
13. Wang, X.; Wei, H. Cryptanalysis of compressive interference-based optical encryption using a U-net deep learning network. *Opt. Commun.* **2022**, *507*, 127641. [CrossRef]
14. Chen, J.; Li, X.-W.; Wang, Q.-H. Deep Learning for Improving the Robustness of Image Encryption. *IEEE Access* **2019**, *7*, 181083–181091. [CrossRef]
15. Jin, M.; Wang, W.; Wang, X. Optical color image cryptosystem based on interference principle and deep learning. *Optik* **2021**, *251*, 168474. [CrossRef]
16. Song, W.; Liao, X.; Weng, D.; Zheng, Y.; Liu, Y.; Wang, Y. Cryptanalysis of phase information based on a double random-phase encryption method. *Opt. Commun.* **2021**, *497*, 127172. [CrossRef]
17. Wang, X.; Wang, W.; Wei, H.; Xu, B.; Dai, C. Holographic and speckle encryption using deep learning. *Opt. Lett.* **2021**, *46*, 5794–5797. [CrossRef]
18. Li, Q.; Meng, X.; Yin, Y.; Wu, H. A Multi-Image Encryption Based on Sinusoidal Coding Frequency Multiplexing and Deep Learning. *Sensors* **2021**, *21*, 6178. [CrossRef]
19. Bao, Z.; Xue, R. Research on the avalanche effect of image encryption based on the Cycle-GAN. *Appl. Opt.* **2021**, *60*, 5320–5334. [CrossRef]
20. Zhao, T.; Ran, Q.; Chi, Y. Image encryption based on nonlinear encryption system and public-key cryptography. *Opt. Commun.* **2015**, *338*, 64–72. [CrossRef]
21. Guan, Z.-H.; Huang, F.; Guan, W. Chaos-based image encryption algorithm. *Phys. Lett. A* **2005**, *346*, 153–157. [CrossRef]
22. Kumar, C.M.; Vidhya, R.; Brindha, M. An efficient chaos based image encryption algorithm using enhanced thorp shuffle and chaotic convolution function. *Appl. Intell.* **2021**, *52*, 2556–2585. [CrossRef]
23. Liu, X.; Tong, X.; Wang, Z.; Zhang, M. A new n-dimensional conservative chaos based on Generalized Hamiltonian System and its' applications in image encryption. *Chaos Solitons Fractals* **2022**, *154*, 111693. [CrossRef]
24. Yang, Y.G.; Wang, B.P.; Yang, Y.L.; Zhou, Y.H.; Shi, W.M.; Liao, X. A visually meaningful image encryption algorithm based on adaptive 2D compressive sensing and chaotic system. *Multimed. Tools Appl.* **2022**, 1–30. [CrossRef]
25. Chai, X.; Wu, H.; Gan, Z.; Han, D.; Zhang, Y.; Chen, Y. An efficient approach for encrypting double color images into a visually meaningful cipher image using 2D compressive sensing. *Inf. Sci.* **2021**, *556*, 305–340. [CrossRef]
26. Kiran, P.; Parameshachari, B.D. Logistic Sine Map (LSM) Based Partial Image Encryption. In Proceedings of the 2021 National Computing Colleges Conference (NCCC), Taif, Saudi Arabia, 27–28 March 2021; pp. 1–6. [CrossRef]
27. Ali Khan, F.; Ahmed, J.; Ahmad, J.; Khan, J.S.; Ahmad, F.; Stankovic, V.; Larijani, H. A novel chaos-based partial image encryption scheme using Lifting Wavelet Transform. In Proceedings of the 1st International Nonlinear Dynamics Conference, Rome, Italy, 17–20 February 2019.
28. Bhowmik, A.; Karforma, S.; Dey, J.; Sarkar, A. Fuzzy-Based Session Key as Restorative Power of Symmetric Key En-cryption for Secured Wireless Communication. In Proceedings of the 2nd International Conference on Communication, Devices and Computing. Lecture Notes in Electrical Engineering; Kundu, S., Acharya, U., De, C., Mukherjee, S., Eds.; Springer: Singapore, 2020; Volume 602.
29. Bokhari, M.U.; Shallal, Q.M. A review on symmetric key encryption techniques in cryptography. *Int. J. Comput. Appl.* **2016**, *147*, 1504–1518.
30. Padhiar, S.; Mori, K.H. A Comparative Study on Symmetric and Asymmetric Key Encryption Techniques. In *Implementing Data Analytics and Architectures for Next Generation Wireless Communications*; IGI Global: Hershey, PA, USA, 2022; pp. 132–144.
31. Kaur, M.; Kumar, V. A Comprehensive Review on Image Encryption Techniques. *Arch. Comput. Methods Eng.* **2020**, *27*, 15–43. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.