

# DeepSLM: Speckle-Licensed Modulation via Deep Adversarial Learning for Authorized Optical Encryption and Decryption

Haofan Huang, Qi Zhao, Huanhao Li, Yuandong Zheng, Zhipeng Yu, Tianting Zhong, Shengfu Cheng, Chi Man Woo, Yi Gao, Honglin Liu, Yuanjin Zheng, Jie Tian,\* and Puxiang Lai\*

Optical encryption is pivotal in information security, offering parallel processing, speed, and robust security. The simplicity and compatibility of speckle-based cryptosystems have garnered considerable attention. Yet, the predictable statistical distribution of speckle optical fields' characteristics can invite statistical attacks, undermining these encryption methods. The proposed solution, a deep adversarial learning-based speckle modulation network (DeepSLM), disrupts the strong intercorrelation of speckle grains. Utilizing the unique encoding properties of speckle patterns, DeepSLM facilitates license editing within the modulation phase, pioneering a layered authentication encryption system. Our empirical studies confirm DeepSLM's superior performance on key metrics. Notably, the testing dataset reveals an average Pearson correlation coefficient above 0.97 between decrypted images and their original counterparts for intricate subjects like human faces, attesting to the method's high fidelity. This innovation marries adjustable modification, optical encryption, and deep learning to enforce tiered data access control, charting new paths for creating user-specific access protocols.


## 1. Introduction

We are witnessing an unprecedented surge in data proliferation. The digital universe is expanding exponentially, fueled by the voracious data collection practices of both private entities and governmental bodies.<sup>[1]</sup> Consider Facebook, the behemoth of social networks, which has accumulated over 300 petabytes of user data since its inception<sup>[2]</sup>—a volume that dwarfs the Library of Congress's two centuries' worth of holdings by a factor of one hundred.<sup>[3]</sup> In this era of Big Data, the relentless aggregation and scrutiny of information present formidable challenges to privacy preservation.<sup>[4]</sup> This is particularly critical for biometric identifiers such as facial features, fingerprints, and iris patterns,<sup>[5]</sup> which are imbued with unique and sensitive attributes that are integral to personal identity.

H. Huang, Q. Zhao, H. Li, Y. Zheng, Z. Yu, T. Zhong, S. Cheng, C. M. Woo, P. Lai  
Department of Biomedical Engineering  
Hong Kong Polytechnic University  
Hung Hom, Hong Kong SAR  
E-mail: puxiang.lai@polyu.edu.hk

H. Huang, Q. Zhao, H. Li, Y. Zheng, Z. Yu, T. Zhong, S. Cheng, C. M. Woo, H. Liu, P. Lai  
Shenzhen Research Institute  
Hong Kong Polytechnic University  
Shenzhen 518057, China

Y. Gao  
School of Biomedical Engineering  
Shenzhen University Medical School  
Shenzhen University  
Shenzhen 518055, China

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202400150>.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202400150

H. Liu  
Key Laboratory for Quantum Optics  
Shanghai Institute of Optics and Fine Mechanics  
Chinese Academy of Sciences  
Shanghai 201800, China

Y. Zheng  
School of Electrical and Electronic Engineering  
Nanyang Technological University  
Singapore 639798, Singapore

J. Tian  
Beijing Advanced Innovation Center for Big Data-Based Precision Medicine  
School of Medical Science and Engineering  
Beihang University  
Beijing 100191, China  
E-mail: tian@ieee.org

J. Tian  
Key Laboratory of Molecular Imaging  
Institute of Automation  
Chinese Academy of Sciences  
Beijing 100190, China

For an adult, altering these biometric markers is typically infeasible, rendering individuals highly vulnerable to identity theft in the event of a data breach.<sup>[6]</sup> The spate of database compromises at major corporations underscores the pressing imperative for the development and implementation of robust security measures to protect sensitive information from illicit access.<sup>[7]</sup>

Optical encryption harnesses the diverse properties of light, such as phase, amplitude, wavelength, polarization, and orbital angular momentum, to encode and decode vital information. This approach heralds a new frontier for secure communication and data storage.<sup>[8–10]</sup> Unlike traditional digital encryption, which depends on digital signal processing, optical encryption leverages the inherent multiplexing capabilities and multiple degrees of freedom of light. These features facilitate parallel processing, high-speed operations, and enhanced security.<sup>[7,9,11,12]</sup> Since the advent of the dual-random phase encoding technique by Refregier et al.<sup>[13]</sup> there has been a significant evolution in optical image encryption. This progress has spawned a variety of cryptographic systems, including diffraction imaging,<sup>[14]</sup> digital holographic encryption,<sup>[15]</sup> phase retrieval algorithms,<sup>[16]</sup> compressed sensing,<sup>[17]</sup> quantum cryptography,<sup>[18]</sup> and chaotic systems.<sup>[19]</sup> However, the complexity of optical designs and the sophisticated techniques required for phase information retrieval have posed challenges to system integration, impeding broader practical application.

In recent years, with the integration of deep learning into optical information processing, attention has been shifted toward optical speckle encryption schemes based on deep learning. Pioneering work by Chen et al. showcased the first deep learning-based solution in optical encryption, utilizing neural networks to extract plaintext information from speckle patterns.<sup>[20]</sup> Following this, Dai et al. introduced a holographic speckle encryption method employing dense convolutional neural networks,<sup>[21]</sup> while Wang et al. developed a technique for encrypting phase images into speckle patterns via random amplitude modulation during optical transmission.<sup>[22]</sup> The decryption process involves reconstructing the original image from the ciphertext using an advanced U-Net model. Additional noteworthy contributions include Zhou et al. proposed a learning-based optical verification method in complex scattering media,<sup>[23]</sup> Feng et al.'s speckle encryption and transmission technique leveraging orbital angular momentum,<sup>[24]</sup> and Zhao et al.'s application of speckles in encrypted face recognition with extended key lengths.<sup>[25]</sup> The simplicity of the optical setups and their compatibility with other systems make speckle-based encryption an optimal choice for safeguarding optical information.

The potential of speckle-based encryption in optical communication is vast, yet it necessitates further exploration to validate its security measures. The intensity of the speckle light field typically adheres to a negative exponential distribution, as illustrated

in **Figure 1b**, which contrasts with the ideal ciphertext's statistical profile depicted in **Figure 1d**.<sup>[26]</sup> To thwart statistical attacks, encrypted image pixels must exhibit a uniform or near-uniform distribution.<sup>[27,28]</sup> Traditional speckle-based encryption methods may inadvertently disclose statistical data about the plaintext during encryption, allowing adversaries to deduce information by scrutinizing the ciphertext's statistical attributes. Consequently, encryption algorithm design must prioritize the statistical characteristics of ciphertext to bolster data confidentiality and integrity.<sup>[29]</sup>

Moreover, speckle grains within a pattern are not entirely autonomous but exhibit physical correlations over various ranges, unlike the negligible dependencies among pixels in an ideal ciphertext.<sup>[30,31]</sup> Breaking these correlations is essential for securing information. Additionally, the prevailing centralized model for consent in information permission challenges the principle of user control, a fundamental aspect of Fair Information Practices, restricting users' capacity for nuanced decision-making regarding third-party data access.<sup>[4,32]</sup> Hence, there is a pressing need for an authorized access framework that empowers users to govern the disclosure and usage of their identity data.

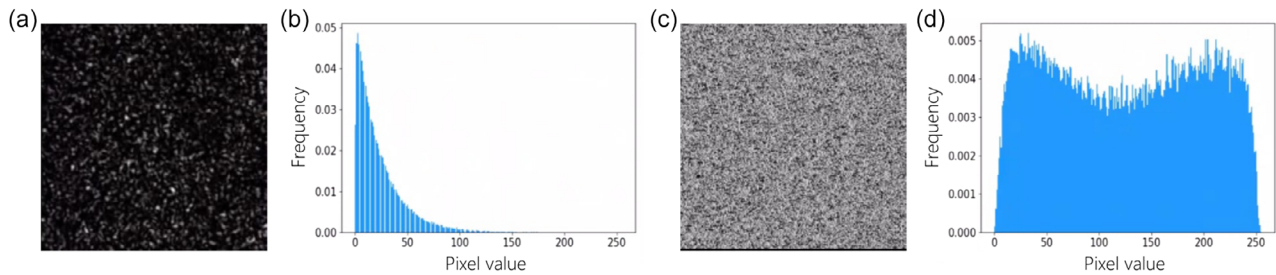
The challenge lies in integrating permission information into the encryption and decryption phases to regulate access. For instance, license-modulated encryption and authorized decryption could enable high-priority users to view clear plaintext, while low-priority users might only access blurred or watermarked versions. The nature of multiple light scattering and the virtually limitless transmission channels in scattering media disperses plaintext information into a particle-like state throughout the speckle pattern's field of view.<sup>[33–35]</sup> This distinctive one-to-many mapping permits the inclusion of extra information without jeopardizing the retrieval of the original content.<sup>[36,37]</sup> Inspired by this, we propose modulating the speckle during encryption to create a tiered authentication system for information access.

Based on these observations, in this work, we propose a novel deep adversarial learning-based speckle modulation network (DeepSLM) for license modulation of speckle patterns, facilitating authorized optical encryption and decryption. **Figure 2** shows an overview of the proposed system, which encompasses two core processes: modulated encryption and authorized decryption. During the encryption phase, the optical setup first transcribes the plaintext into a speckle pattern. Subsequently, DeepSLM's encryption subnetwork applies licensed modulation, producing ciphertext that boasts augmented randomness and embedded authorization checks. In the decryption stage, the level of authorization dictates the extent of information available to external parties. High-priority individuals are granted access to the unaltered plaintext, whereas low-priority individuals are restricted to viewing only the blurred or watermarked versions of the plaintext.

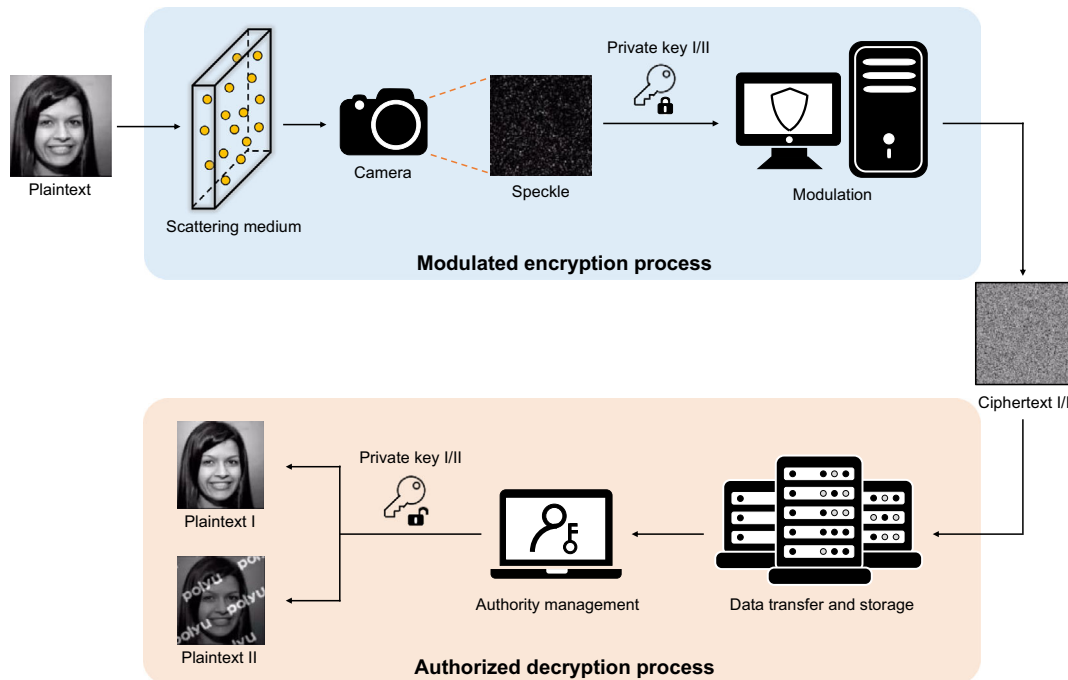
Our system adeptly conceals plaintext information during the data collection phase. The resulting speckle pattern, once modulated and encrypted, can be securely stored and transmitted, effectively shielding private data from potential leaks. Upon decryption, specific authorization levels are necessary to retrieve the corresponding plaintext, thereby enforcing a tiered data access hierarchy. Our analysis reveals that DeepSLM has effectively disrupted the intrinsic correlation among speckle particles, yielding ciphertext with superior randomness and security. Furthermore, the average Pearson correlation coefficient

P. Lai  
Photonics Research Institute  
Hong Kong Polytechnic University  
Hung Hom, Hong Kong SAR

P. Lai  
Research Institute for Sports Science and Technology  
Hong Kong Polytechnic University  
Hung Hom, Hong Kong SAR



**Figure 1.** a,b) An example of a speckle pattern and its histogram. c,d) An example of ideal ciphertexts and its histogram.



**Figure 2.** The framework of the proposed authorized speckle-based encryption and decryption system. Modulated encryption: the plaintext image is encrypted into a speckle pattern, which is then modulated by DeepSLM to enhance randomness and integrate authorization. Authorized decryption: different authorization levels determine the decrypted images with/without watermarks.

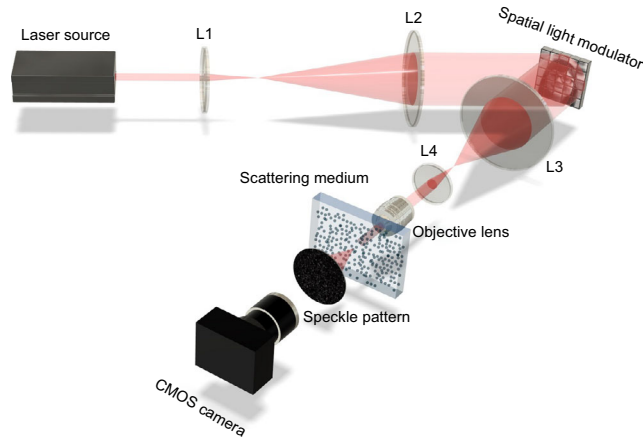
(PCC) between the decrypted and original images in our test dataset surpasses 0.97, signifying exceptional information fidelity of our method and its promise for a wide array of applications. Security assessments confirm the robustness of our method against diverse attack strategies. In tests of randomness, our method demonstrates leading-edge performance, excelling in most evaluation metrics. Overall, our system offers robust protection for biometric templates, such as original face images, against security threats, while also facilitating their flexible replacement, which has significant implication for privacy preservation in the Big Data landscape.

## 2. Methodology

### 2.1. Data Acquisition

The image information is encoded into a speckle pattern through the designed optical system. **Figure 3** illustrates the optical setup

used in encryption experiments. In this work, face images from the “Flickr Faces High Quality” (FFHQ) database<sup>[25,38–40]</sup> are used. First, we modulate the incident coherent light from a 532 nm single mode laser source (EXLSR-532-300-CDRH, Spectra-Physics, USA) by projecting the images on a phase-modulating Spatial Light Modulator (SLM, HOLOEYE PLUTO VIS056 1080p, German). Thus, the modulated laser beam can carry the information of the face images. The modulated wave fronts are then scattered through a scattering medium (220-grid ground glass, DG10-220-MD, Thorlabs, USA) to yield random speckles recorded by a CMOS camera (FL3-U3-32S2M-CS, PointGrey, Canada). In order to confirm that each speckle pattern is matched to the face image projected on the spatial light modulator, we employ MATLAB software to synchronize all the devices during experiments: Initially, plaintext is transferred to the SLM and displayed. This is achieved through MATLAB commands that interface with the SLM’s controlling software. After the plaintext is displayed on the SLM, we introduce a brief



**Figure 3.** The optical setup for information encryption. Face images are loaded on the spatial light modulator and are illuminated by a 532 nm continuous coherent laser source. The modulated laser beam passes through the scattering medium, resulting in optical speckles captured by a CMOS camera.

pause of 0.05 s within the MATLAB script. This pause allows the SLM to stabilize the display of the plaintext image, ensuring consistency in the speckle patterns generated. Once the plaintext is stably displayed, the camera is triggered via MATLAB to capture the corresponding speckle pattern. This step is critical for matching each speckle pattern with the projected face image.

## 2.2. Authorized Encryption and Decryption Network Design

At this stage, a permission-tunable encryption and decryption network based on deep learning speckle-licensed modulation is proposed. The DeepSLM is illustrated in **Figure 4**, which mainly consists of three subnetworks: 1) the license modulation network  $G$ , 2) the projection discriminator network  $D$ , and 3) the authorized decryption network  $F$ . Given a speckle pattern  $x$  and licensed label  $c$ , they are first fed into the license modulation network  $G$  to generate the ciphertext with permission information. Notably, at this stage, different domain labels lead to different modulated ciphertexts: those modulated by licensed label  $c_1$  are related to the high-definition face image (first-level authority), and those modulated by licensed label  $c_2$  define the watermarked image (second-level authority). Then, the authorized decryption network  $F$  transforms the ciphertext and the corresponding licensed label to the plaintext with/without watermark. The projection discriminator network  $D$  is mainly designed to enhance the encryption performance of the modulation network under the configuration of the generative adversarial network (GAN) training.

### 2.2.1. License Modulation Network $G$

The architecture of the network  $G$ , as shown in **Figure 4**, combines an encoder, a decoder, and six residual blocks. The encoder is composed of convolutional layers that adopt a stride of two for downsampling, while the decoder is composed of deconvolution layers with a stride of two for upsampling. Our goal is to train a

single generator  $G$  that translates an input image  $x$  (i.e., speckle pattern) into an output image  $y$  (i.e., ciphertext) conditioned on the target licensed label  $c$ , i.e.,  $G(x, c) \rightarrow y$ . The key idea is to switch the affine scaling parameters of the batch normalization (BN)<sup>[41]</sup> in the generator  $G$  by embedding the licensed label  $c$ , which is called conditional batch normalization (CBN).<sup>[42]</sup> CBN allows for different data categories to undergo distinct normalization, scaling, and biasing, based on their mean and variance values. Consequently, we no longer rely on the statistical features of the entire training set. Instead, each image is normalized within its feature map, reducing data homogeneity and allowing for license-specific ciphertext modulation. Therefore, CBN is a powerful yet computationally efficient method for modulating neural activations and the network can manipulate feature maps through CBN to encode authority information.

Given a mini-batch  $\mathcal{B} = \{F_{i,\dots}\}_{i=1}^N$  of  $N$  examples, BN normalizes the feature maps during training as follows:

$$BN(F_{i,c,h,w}|\gamma_c, \beta_c) = \gamma_c \frac{F_{i,c,h,w} - E_{\mathcal{B}}[F_{\cdot,c,\dots}]}{\sqrt{\text{Var}_{\mathcal{B}}[F_{\cdot,c,\dots}] + \epsilon}} + \beta_c \quad (1)$$

where  $\epsilon$  is a constant damping factor for numerical stability, and  $\gamma_c$  and  $\beta_c$  are trainable scalars introduced to keep the representational power of the original network. We focus on a single convolutional layer with BN module  $BN(F_{i,c,h,w}|\gamma_c, \beta_c)$ . We would like to directly predict these affine scaling parameters from our label embedding  $e_q$ . A linear layer is therefore used to predict these parameters from the label embedding  $e_q$  for all feature maps within the layer:

$$\hat{\gamma}_c = \text{Linear}_1(e_q), \hat{\beta}_c = \text{Linear}_2(e_q) \quad (2)$$

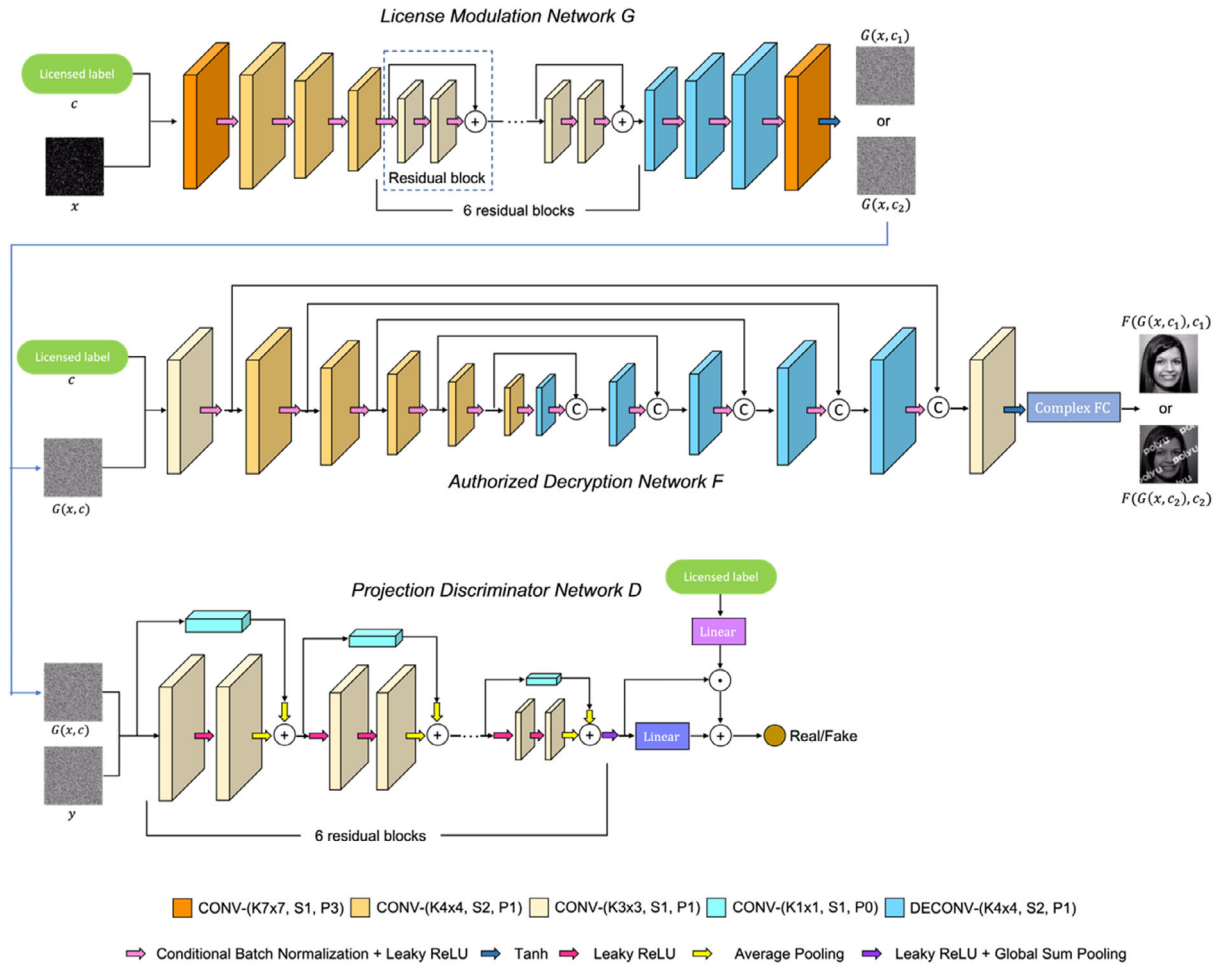
where  $e_q = \sigma(W(c))$ ,  $\sigma$  is the activation function, and  $W$  is the learned parameter weight. Finally, these updated  $\hat{\gamma}_c$  and  $\hat{\beta}_c$  are used as parameters for BN  $BN(F_{i,c,h,w}|\hat{\gamma}_c, \hat{\beta}_c)$  and dependent on the licensed label  $c$ . CBN layers are used in all layers except the last output layer to modulate neural activations. With CBN, the prior inputs  $x$  and  $c$  are combined in a joint hidden representation, which facilitates the network to generate feature maps of different permissions.

Conditional on the licensed label  $c$ , the generator  $G$  is enabled to modulate different permission information. Here, the label  $c$  can be treated as any type of supplementary information, such as category labels or other modality data, which depends on its application. It extends GANs to a conditional model. Given a speckle pattern  $x$  and the licensed label  $c$  as inputs, the generator learns a mapping function from the speckle distribution  $p_x$  to the ciphertext data space as  $G(x, c)$ . The loss  $L_G$  of the network  $G$  is

$$L_G = -\mathbb{E}_{x \sim p_x} D(G(x, c), c) \quad (3)$$

The aim of  $L_G$  is to minimize the accuracy of the network in identifying the ciphertext produced by the network  $G$ . In this work, we select the category labels as the licensed label and modulate the speckle with two permissions, where permission level I allows the recovery of high-definition images, and permission levels II will embed customized watermark information in the recovered images.





**Figure 4.** The architecture of the proposed encryption and decryption in DeepSLM—license modulation network  $G$ : generate the ciphertext with permission information; authorized decryption network  $F$ : decrypt the ciphertext according to the corresponding licensed label; projection discriminator network  $D$ : enhance the encryption performance of the modulation network.

### 2.2.2. Projection Discriminator Network $D$

The role of the network  $D$  is to control the style of ciphertext generation. As mentioned above (see Figure 1), the speckle pattern does not satisfy the chaotic distribution, so it is necessary to transform the output image distribution to strengthen the security of the ciphertext. We adopt the method proposed by Ding et al.<sup>[43]</sup> and set the target domain as a chaotic style with a high level of security. The network  $D$  will evaluate the image generated by the network  $G$  to determine whether it belongs to the target domain. The feature extraction part of network  $D$  consists of six residual blocks. Each residual block uses two convolutional layers to extract features of the input image, and finally the resultant high-level hidden feature vectors are used in the final discriminative module. The network  $D$  outputs a single scalar that reflects the probability of that the input image came from the target domain  $p_{\text{data}}$  as opposed to the source domain  $p_x$ . The loss function  $L_D$  is as follows:

$$L_D = \mathbb{E}_{y \sim p_{\text{data}}} [\max(0, 1 - D(y, c))] + \mathbb{E}_{x \sim p_x} [\max(0, 1 + D(G(x, c), c))] \quad (4)$$

which follows the hinge version of the standard adversarial loss.<sup>[44–46]</sup>  $x$  represents the speckle pattern, and  $y$  represents the data from the target transformation domain.

To stabilize the training of GAN and improve the quality of the generator, we employ a projection-based approach to incorporate conditional information into the discriminator. According to Miyato et al.<sup>[46]</sup> the optimal discriminator for conditional GAN is described as follows:

$$D(z, c; \theta) := f_1(z, c; \theta) + f_2(z; \theta) = c^T V \phi(z; \theta_\phi) + \psi(\phi(z; \theta_\phi); \theta_\psi) \quad (5)$$

where  $z$  is the input vector that encompasses both the images synthesized by network  $G$  and the target domain data  $y$ ,  $\theta$  is the parameters of  $D$ ,  $f_1$  and  $f_2$  are some parametric functions,  $V$  is the embedding matrix of  $c$ ,  $\phi(\cdot, \theta_\phi)$  is a vector output function of  $z$ , and  $\psi(\cdot, \theta_\psi)$  is a scalar function of the same  $\phi(z; \theta_\phi)$ . The discriminator structure produced by Equation (5) requires us to perform an inner product between the feature vector and the embedding condition vector, as shown in Figure 4. We encode conditional information into conditional vectors by using a linear

layer, and the learned parameters  $\theta = \{V, \theta_\Phi, \theta_\Psi\}$  are to be trained to optimize the adversarial loss defined by Equation (4).

### 2.2.3. Authorized Decryption Network $F$

The function of the network  $F$  is to recover the plaintext information from the ciphertext that is generated by the network  $G$ . As shown in Figure 4, the design of the network  $F$  is based on the widely applied U-Net architecture,<sup>[47]</sup> equipped with CBN layers and a complex fully connected layer.<sup>[48]</sup> The encoder of the network  $F$  includes the convolutional layers with a stride of two to downsample the input data and extract the hidden feature vector. The decoder of the network  $F$  uses the deconvolution layers to upsample the feature vector and then extracts high-dimensional feature representations to reconstruct the image. The CBN layer is used to integrate conditional information into the decryption process, whose parameters are the corresponding decryption keys. Finally, the network extracts features of different dimensions through the encoder and decoder, adjusts the feature map through the CBN layer to control the decryption authority, and outputs the plaintext corresponding to the security level. The reconstruction loss  $L_F$  used for training the network  $F$  is defined as

$$L_F = MSE(\hat{g}, g) - PCC(\hat{g}, g) \\ = MSE(F(G(x, c), c), g) - PCC(F(G(x, c), c), g) \quad (6)$$

$$PCC = \frac{\text{mean}[(g - \text{mean}(g)) \times (\hat{g} - \text{mean}(\hat{g}))]}{\text{std}(g) \times \text{std}(\hat{g})} \quad (7)$$

$$MSE = \text{mean}[(\hat{g} - g)^2] \quad (8)$$

where  $g$  and  $\hat{g}$  are the ground truth and the decrypted images by network  $F$ , respectively.

## 2.3. Implementation Details

20 000 pairs of speckle patterns and human face images were selected for the initial DeepSLM training: 19 900 image–speckle pairs for training and 100 image–speckle pairs to evaluate the performance and applicability of the network. The face images were obtained from the FFHQ dataset, and the custom watermark was added using OpenCV in Python. The speckle patterns (Figure 3) were obtained from the optical setup. During training, the resolution of the input speckle pattern and the output modulated ciphertext was set to  $256 \times 256$ , and the resolution of the output decrypted image was  $64 \times 64$ .<sup>[25]</sup> We utilized the Adam optimizer<sup>[49]</sup> to optimize the DeepSLM. The exponential decay rate of the first-order moment estimation was 0.5, and the second-order moment estimation was 0.999. The network's weight parameters were initialized randomly, and a batch size of 24 was selected. The learning rate was set to 0.0001 initially and the training epoch was configured to 50 000 iterations to achieve better performance. Our whole algorithm was implemented based on PyTorch,<sup>[50]</sup> which is installed on a PC with Intel(R) Xeon(R) Gold 5218R CPU @ 2.10 GHz and NVIDIA RTX A6000 GPU.

## 3. Results and Discussions

### 3.1. Performance of Encryption and Decryption

The evaluation of the encryption and decryption performance is conducted on the testing data set, and the experimental results are shown in Figure 5. Plaintext I is the original image, and plaintext II is the original image with a custom watermark added. By feeding the plaintext I into the optical system shown in Figure 3, the original image is encoded as an optical speckle. Then, the proposed encryption network can modulate the speckle into different levels of ciphertext (ciphertext I/II) according to different authority keys. Correspondingly, the decryption network needs to use a key that matches the encryption level to correctly decrypt the ciphertext. The decryption plaintexts I and II are the result of the decryption of ciphertexts I and II by the decryption network, respectively. Evidently, the ciphertext image produced by the encrypted network is notably distinct from the original one, showing no discernible identity information. At the same time, the decryption network can effectively restore the image with high fidelity. In addition, we conduct histogram analysis, entropy analysis, and correlation analysis on the ciphertext image to further verify the effectiveness of encryption. Four metrics are used, PCC, mean square error (MSE), structural similarity index measure (SSIM), and peak signal-to-noise ratio (PSNR), to measure the similarity between the decrypted image and the original image.

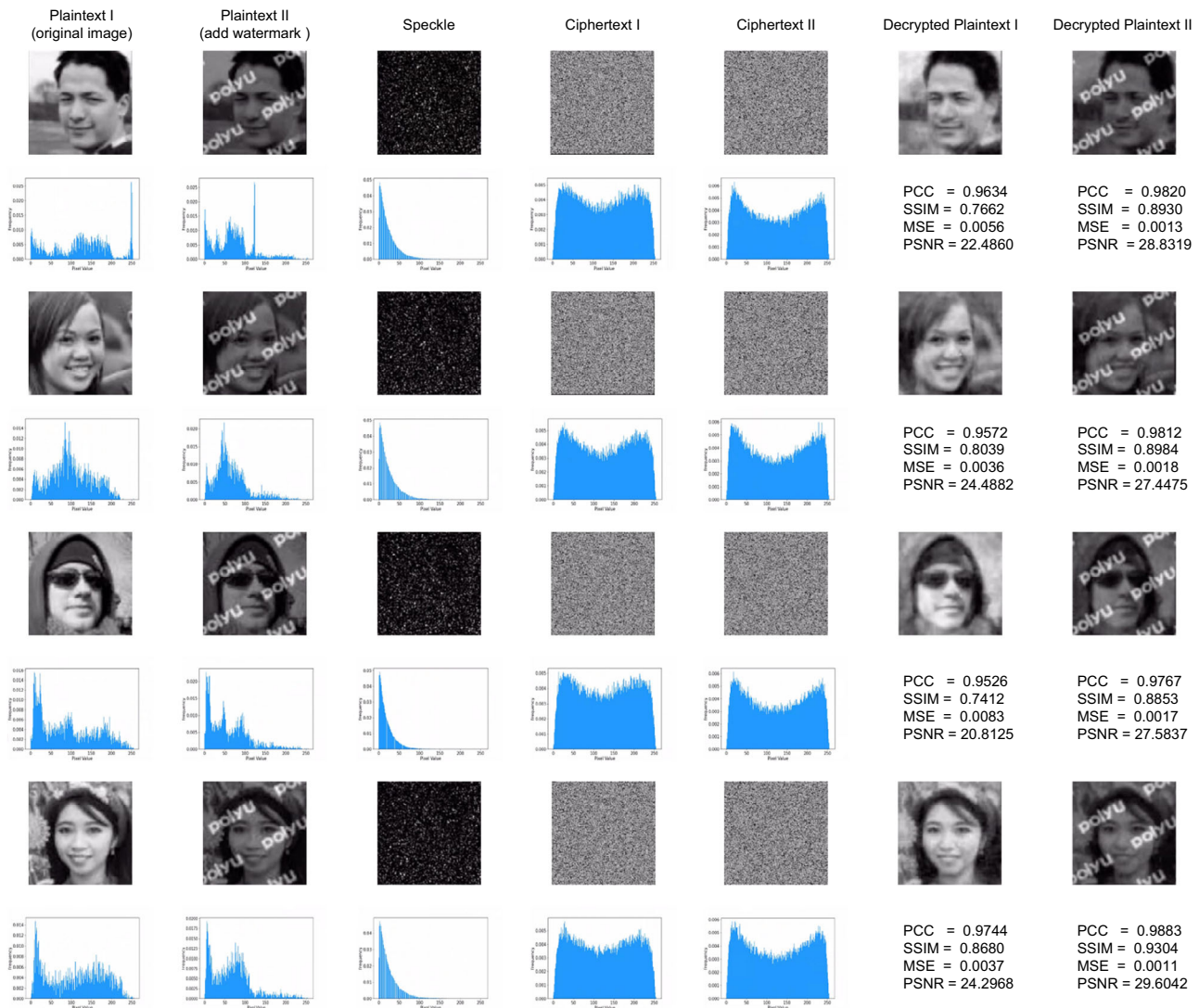
#### 3.1.1. Encryption Performance Analysis

**Histogram Analysis:** The histogram provides a clear and intuitive image statistical representation by displaying the distribution of gray levels in the image.<sup>[51]</sup> The variance of the histogram is an effective measure to evaluate the ability of an encryption algorithm to withstand statistical analysis attacks. The smaller the variance, the more uniform the pixel distribution, the less statistical information the image displays, and the safer the image encryption scheme. As shown in Figure 5, the histogram of the plaintext image and speckle show obvious statistical rules, and the distribution of ciphertext images is more uniform, increasing its immunity to statistical attacks.

**Entropy Analysis:** Information entropy is another important reference index to measure the randomness of information.<sup>[51]</sup> The information entropy of an image is a statistical method of assessing random features, which can serve as a measure for the random properties of images. It calculates the pixels' distribution for each gray level in every color channel, and a more uniform distribution indicates a higher level of resistance to statistical attacks. The information entropy is defined as follows:

$$H(X) = - \sum_{i=1}^L P(X_i) \log_2 P(X_i) \quad (9)$$

where  $P(X_i)$  is the probability that the pixel value  $X_i$  appears and  $L$  is the number of gray levels. Table 1 shows the results calculated according to Equation (9). The average information



**Figure 5.** Examples of encrypted ciphertexts and decrypted plaintexts in the proposed cryptosystem. Intensity histograms are labeled under plaintext I, plaintext II, speckle, ciphertext I, and ciphertext II. Similarities between decrypted plaintext images and their corresponding ground truth images are marked under decrypted plaintext I and decrypted plaintext II.

**Table 1.** Entropy and correlation analysis of the ciphertext from the proposed license modulation network. The bold values in the table represent the best performance under the current evaluation metric.

Metrics	Entropy	Correlation		
		Horizontal	Vertical	Diagonal
Plaintext	6.4197 ± 0.3213	0.9649 ± 0.0108	0.9647 ± 0.0149	0.9431 ± 0.0163
Speckle	5.6901 ± 0.0075	0.7942 ± 0.0230	0.7776 ± 0.0138	0.6338 ± 0.0219
Ciphertext	<b>7.9536 ± 0.0037</b>	<b>0.0186 ± 0.0094</b>	<b>0.0254 ± 0.0176</b>	<b>0.0180 ± 0.0156</b>

entropy of our encrypted images is 7.9536, approaching the theoretical maximum of 8. This high entropy level signifies substantial uncertainty in the image content, indicating effective resistance to statistical attacks.

**Correlation Analysis:** The correlation of adjacent pixels indicates the level of correlation of pixel values in neighboring areas of the image. Adjacent pixels in a digital image show high correlation, and a pixel often leaks information about its surrounding pixels. This characteristic can be exploited by the attacker to deduce the gray value of adjacent pixels, thereby realizing the restoration of the entire plaintext image.<sup>[51]</sup> Because of this, this strong correlation must be broken. In experiment, 1000 pairs of adjacent pixels were randomly selected from images to calculate the correlation in horizontal, vertical, and diagonal directions. The correlation coefficient between the two adjacent pixels was calculated according to Equation (7). As shown in Table 1, for the plaintext image and speckle, the correlation between adjacent pixels is strong, while that of the ciphertext image is extremely weak. This indicates that the encryption method can severely weaken the correlation of adjacent pixels.

### 3.1.2. Decryption Performance Analysis

During the testing, PCC, MSE, SSIM, and PSNR were employed as the criteria for measuring image similarity, as defined in Equation (7), (8), (10), and (11) respectively:

$$SSIM(\hat{g}, g) = \frac{(2\mu_{\hat{g}}\mu_g + C_1)(2\sigma_{\hat{g}g} + C_2)}{(\mu_{\hat{g}}^2 + \mu_g^2 + C_1)(\sigma_{\hat{g}}^2 + \sigma_g^2 + C_2)} \quad (10)$$

where  $\mu_{\hat{g}}$  and  $\mu_g$  are the mean value of  $\hat{g}$  and  $g$ , respectively,  $\sigma_{\hat{g}}^2$  and  $\sigma_g^2$  are the variance of  $\hat{g}$  and  $g$ , respectively,  $\sigma_{\hat{g}g}$  is the covariance of  $\hat{g}$  and  $g$ , and  $C_1$  and  $C_2$  are the very small constant ( $10^{-5}$ ) used to maintain the stability. The value range of SSIM is from 0 to 1 and the higher value indicates the higher similarity of the two images. The PSNR is defined as

$$PSNR = 10 \times \log_{10} \frac{(2^n - 1)^2}{MSE} \quad (11)$$

where  $n$  is the number of bits per pixel. For grayscale images, it is generally set as 8, i.e., the number of pixel gray levels is 256. A higher PSNR value indicates a lower degree of image distortion. Four groups of test images are illustrated in Figure 5. The PCC, SSIM, MSE, and PSNR between the original plaintext and the decrypted image are marked below the decrypted images. **Table 2** shows the decryption performance of the DeepSLM on the whole test dataset. The average PCC, MSE, SSIM and PSNR of all test data are 0.9562/0.9818 (image I/II), 0.0056/0.0015, 0.7583/0.8808, and 22.6043 dB/28.3293 dB, respectively. In general, the accuracy of information decryption is very high, and it can meet the standard of passing high-precision face recognition tasks, which is very helpful for practical applications.<sup>[25]</sup>

## 3.2. Security Analysis

### 3.2.1. Key Security Analysis

**Key Space:** The size of the key space determines the resistance to exhaustive attacks. In this study, the number of parameters in the CBN layers of DeepSLM is considered as the size of the key space, and the parameters vary depending on the encryption levels. The explicit specifications of the CBN layers in the encryption and decryption network are shown in **Table 3** and **4**, respectively. In the computer, each key or parameter is a floating-point number of 32 bits that ranges between 0 and 1, and can be converted into the decimal form with ten significant digits. As a result, the encryption model has a key space of  $(2^{32})^{26\,880} (\approx 10^{258\,933})$  and the decryption model has a key space of  $(2^{32})^{29\,952} (\approx 10^{288\,526})$ .

**Table 2.** Similarities between the decrypted images from DeepSLM and the ground truth images.

	Decryption image I	Decryption image II
PCC	0.9562 ± 0.0156	0.9818 ± 0.0037
MSE	0.0056 ± 0.0011	0.0015 ± 0.0002
SSIM	0.7583 ± 0.0422	0.8808 ± 0.0190
PSNR [dB]	22.6043 ± 0.8575	28.3293 ± 0.5936

**Table 3.** Details of CBN layers in the license modulation network G.

CBN layer name	Number	Shape	Parameters	Total parameters
CBN1	1	(64, 6)	384	384
Down CBN1	1	(128, 6)	768	1152
Down CBN2	1	(256, 6)	1536	2688
Down CBN3	1	(512, 6)	3072	5760
Residual block CBNs	6	(512, 6)	18 432	24 192
Up CBN1	1	(256, 6)	1536	25 728
Up CBN2	1	(128, 6)	768	26 496
Up CBN3	1	(64, 6)	384	26 880

**Table 4.** Details of CBN layers in the authorized decryption network F.

CBN layer name	Number	Shape	Parameters	Total parameters
CBN1	1	(64, 6)	384	384
Down CBN1	1	(128, 6)	768	1152
Down CBN2	1	(256, 6)	1536	2688
Down CBN3	1	(512, 6)	3072	5760
Down CBN4	1	(1024, 6)	6144	11 904
Down CBN5	1	(1024, 6)	6144	18 048
Up CBN1	1	(1024, 6)	6144	24 192
Up CBN2	1	(512, 6)	3072	27 264
Up CBN3	1	(256, 6)	1536	28 800
Up CBN4	1	(128, 6)	768	29 568
Up CBN5	1	(64, 6)	384	29 952

Such a large key space provides resistance against exhaustive attacks, as it substantially increases the difficulty for attackers to accurately predict the private key.

**Key Randomness Analysis:** As the training process of the deep learning network is highly stochastic, the private key (network parameters) generated by every training will be different and have a high degree of randomness. During the evaluation, we trained the encryption network 4 times using the same settings, and then we got eight encryption keys, i.e., key A(I/II), key B(I/II), key C(I/II), and key D(I/II), respectively. Then, using the same speckle pattern as the input for the encryption networks with different keys, eight ciphertexts encoding permission information are generated. Next, we calculated the SSIM value between these eight encrypted images to evaluate their similarity, and the results can be found in **Table 5**. The SSIM value between different ciphertexts is mostly lower than 0.01, indicating very low similarity. The uniqueness of the proposed DeepSLM for encryption can be demonstrated, as the generated keys vary significantly even when the conditions are set identically, due to the instability of the neural network training.

**Key Sensitivity Analysis:** In a deep learning model, errors will propagate among layers and accumulate as the network gets deeper. In the convolutional layer, for example, the pixel values of the feature map are passed through the convolution kernel to the neighboring pixels in the next layer. Once a parameter error



**Table 5.** SSIM between eight different ciphertexts.

Image <sup>a)</sup>		A		B		C		D	
		I	II	I	II	I	II	I	II
A	I	1	0.0193	0.0039	0.0086	0.0046	0.0022	0.0078	0.0081
	II	0.0193	1	0.0039	0.0058	0.0070	0.0042	0.0067	0.0071
B	I	0.0039	0.0039	1	0.1272	0.0100	0.0078	0.0036	0.0035
	II	0.0086	0.0058	0.1272	1	0.0024	0.0029	0.0062	0.0048
C	I	0.0046	0.0070	0.0100	0.0024	1	0.2296	0.0079	0.0151
	II	0.0022	0.0042	0.0078	0.0029	0.2296	1	0.0114	0.0149
D	I	0.0078	0.0067	0.0036	0.0062	0.0079	0.0114	1	0.1513
	II	0.0081	0.0071	0.0035	0.0048	0.0151	0.0149	0.1513	1

<sup>a)</sup>The images A(I/II), B(I/II), C(I/II), and D(I/II) are the outputs obtained by inputting the same speckle pattern into the encryption networks using different encryption keys.

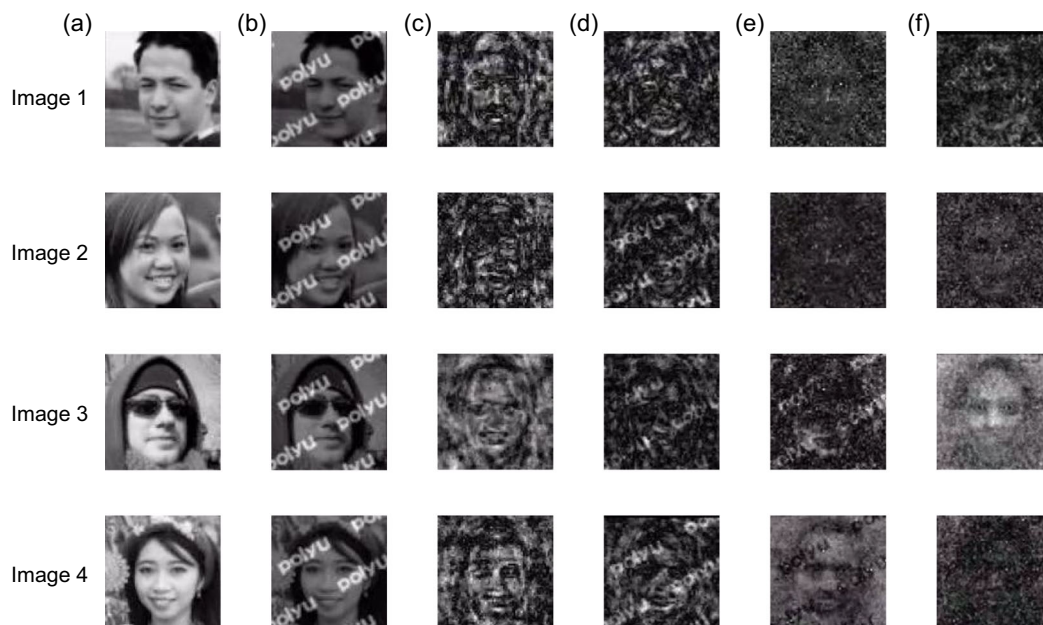
occurs, the error will propagate along the same path and the range of affected pixels will gradually expand as the depth of the network increases. In particular, during upsampling, this error will grow exponentially with the superposition of the deconvolution operations. We presume that the attacker is aware of most private keys (e.g., 95%), with only a small fraction (5%) of key parameters altered and considered as unknown components. Then, the new parameters are fed into the encryption network and the output ciphertext is fed into the original decryption network for information recovery. Similarly, the output ciphertext from the original encryption network is fed into the decryption network with the new parameters to see if the information can be decrypted correctly.

Figure 6 shows the experimental results, where column (a) shows high-definition face images and column (b) shows the images with customized watermark. The images displayed in columns (c) and (d) are the result of decrypting the ciphertext (using the original decryption key) which was encrypted using the modified encryption key. The images in columns (e) and (f) are the result of recovering the original encrypted image utilizing the modified decryption key. It can be seen that whether the parameters of the encryption network or the decryption network are destroyed, the plaintext information cannot be recovered. This implies that even if a small portion (e.g., 5%) of the parameters are altered, the private key is unable to correctly encrypt or decrypt the plaintext information. In other words, it is difficult for a hacker to attack the proposed network by guessing most of the key information in such a large key space.

### 3.2.2. Security Analysis Under Different Attack Models

**Ciphertext Only Attack:** The ciphertext only attack (COA) presumes that the attacker can obtain one or more ciphertexts that have been encrypted using the same key and deduce the plaintext or key by analyzing these ciphertexts.<sup>[52]</sup> However, the encryption model has a key space of up to  $(2^{32})^{26\ 880}$  bits (see Table 3). This denotes the immense difficulty for an attacker to accurately conjecture the private key and recover the plaintext. In addition, our experimental results (see Table 5) show that the key generation process is highly random and the ciphertext image is complex. Therefore, it is challenging to break our encryption model via a COA.

**Known Plaintext Attack:** A known plaintext attack (KPA) assumes that the attacker has access to a portion of the plaintext and its corresponding ciphertext, such as obtaining the front



**Figure 6.** Decryption performance of different keys: a,b) are the face images of different permission levels (level I/II); c) (level I) and d) (level II) are the decryption results of inputting the output ciphertext which obtained from the encryption network with the updated key (change 5% of parameters) to the original decryption network; and e) (level I) and f) (level II) are the outputs of the decryption network after updating the key (change 5% of parameters) by feeding the original ciphertexts.

portion of the data to crack the encryption algorithm, so as to obtain the latter part of the ciphertext.<sup>[52]</sup> According to Table 1, the information entropy of the ciphertext is relatively high and the correlation within the ciphertext is exceedingly low, thereby preventing the attacker from inferring the complete plaintext through contextual analysis of the ciphertext. In addition, it can be clearly seen from Figure 5 that the plaintext and ciphertext are very different, which greatly increases the difficulty of the attack. Consequently, our method resists KPA.

**Chosen Plaintext Attack:** Through this attack method, the attacker can gain access to the encrypted device and thus obtain the corresponding plaintext and ciphertext pairs.<sup>[52]</sup> They typically modify the original plaintext image in a targeted manner and then encrypt the changed image and the original plaintext image separately using the encryption device. Subsequently, the attacker compares these two encrypted ciphertext images to establish the correlation between the plaintext image and the ciphertext image. Finally, the attacker uses this relationship to decipher the ciphertext image. In order to resist chosen plaintext attacks (CPA), when the pixels of the plaintext image are slightly altered, the ciphertext image output by the encryption method must have a substantial change. The larger the variation, the more robust the defense ability against CPA will be. Here, we use the number of pixel change rate (NPCR) and the unified average changing intensity (UACI)<sup>[53]</sup> to measure the pixel-level difference between the ciphertext output by the proposed network when the plaintext image undergoes a small number of pixel changes (only 1% of the pixels change). The NPCR implies the ratio of unequal pixels at the same locations between two images relative to the total number of pixels in the images. The definition of NPCR is as follows:

$$NPCR = \sum_{i,j} \frac{D(i,j)}{T} \times 100\% \quad (12)$$

where

$$D(i,j) = \begin{cases} 0, & \text{if } C^1(i,j) = C^2(i,j) \\ 1, & \text{if } C^1(i,j) \neq C^2(i,j) \end{cases} \quad (13)$$

where  $C^1$  and  $C^2$  represent the ciphertext images before and after a small number of pixel changes in a plaintext image, respectively. The pixel values at grid  $(i,j)$  in  $C^1$  and  $C^2$  are denoted as  $C^1(i,j)$  and  $C^2(i,j)$ . Symbol  $T$  denotes the total number pixels in the ciphertext. UACI denotes the number of averaged changed intensity between ciphertext images. The definition of UACI is as follows:

$$UACI = \sum_{i,j} \frac{|C^1(i,j) - C^2(i,j)|}{F \cdot T} \times 100\% \quad (14)$$

where  $F$  denotes the largest supported pixel value compatible with the ciphertext image format. In this work,  $F$  is set to 255. The results of the experiments are presented in Table 6. It can be seen that a minor alteration to the original image (change only 1% of pixels) can result in a significant difference between two generated ciphertexts (NPCR > 97% and UACI > 48%). The results show that our proposed encryption model can

**Table 6.** The results of plaintext sensitivities to the CPA.

Image ID	1		2		3		4	
Ciphertext ID	I	II	I	II	I	II	I	II
NPCR [%]	97.90	97.80	97.37	97.06	97.53	97.21	97.86	97.66
UACI [%]	48.74	49.15	48.58	48.62	48.84	48.88	48.76	48.98

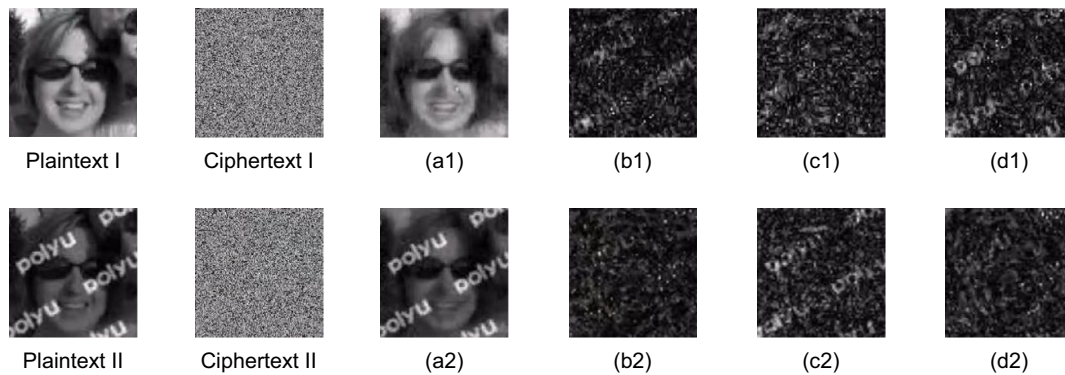
effectively defend against CPA because it is very sensitive when tiny differences occur in plaintext.

**Chosen Ciphertext Attack:** In such attacks, the attacker can arbitrarily create or choose some ciphertexts and get their corresponding plaintexts. By collecting and analyzing these specific ciphertext-plaintext pairs extensively, the key might be cracked successfully.<sup>[52]</sup> As the decryption process also relies on a deep trainable model, the experiment of the chosen ciphertext attack (CCA) is comparable to CPA experiments. For this experiment, the inputs to the decryption network are the original and the modified ciphertext images, respectively. Then the NPCR and UACI are used to compute the difference between the corresponding decrypted images. According to the experimental results (shown in Table 7), when the ciphertext image is slightly modified (just 1% pixels changed), the NPCR value and UACI value between the two decrypted images are more than 98% and 28%, respectively. This demonstrates the superior diffusion features of our decryption model and the fact that a little modification to the ciphertext image can result in a significantly different decrypted image. The proposed model is also very sensitive to ciphertexts and can effectively resist CCA.

**Imitation Learning Attack:** As our method is based on a trainable neural network, we will evaluate whether it is possible for an attacker to employ an imitation learning attack to produce an appropriate key capable of decrypting the target ciphertext image. In this experiment, we assume that the attacker is aware of all our experimental settings, such as network architecture and data processing methods. In other words, we trained four encrypt-decrypt networks (networks A, B, C, D) using the same training conditions. At the initial stage of the training process, the network parameters are initialized randomly. Next, the ciphertext encrypted by the trained network A is input into the decryption networks B, C, and D, respectively. Through evaluating the decryption precision of these networks on a common ciphertext image, the distinctiveness of the parameters generated by the networks can be verified. The results are shown in Figure 7, where the ciphertext encrypted by network A cannot be correctly decrypted by networks B, C, and D. The results demonstrate that even if attackers have knowledge of the network architecture and training conditions, they are still unable to generate an identical

**Table 7.** The results of ciphertext sensitivities to the CCA.

Image ID	1		2		3		4	
Plaintext ID	I	II	I	II	I	II	I	II
NPCR [%]	98.60	98.02	99.41	99.17	99.07	98.54	99.27	98.71
UACI [%]	32.07	28.93	35.14	33.01	33.50	33.85	33.22	31.18



**Figure 7.** Analysis results of imitation learning attacks. The ciphertext image is encrypted by encryption network A and then input to decryption networks A, B, C, and D, respectively. a1–d2) The decrypted results they output are shown, respectively. It can be seen that even with the same model structure and training conditions, networks B, C, and D cannot correctly decrypt images encrypted by network A. This means that the parameters of the networks, i.e., the keys, are all different.

**Table 8.** *P*-value of each encryption algorithm under four metrics.

	Nonoverlapping template matching	Binary matrix rank	Maurer's universal test	Random excursions variant
ECC	$0.9895 \pm 0.0257$	$0.3595 \pm 0.2291$	$<0.01$	<b><math>0.2766 \pm 0.1270</math></b>
RSA	$0.9912 \pm 0.0133$	$0.3349 \pm 0.2625$	$<0.01$	$0.2617 \pm 0.1249$
Speckle	$0.9638 \pm 0.0908$	$0.3670 \pm 0.2354$	$<0.01$	$0.2454 \pm 0.1295$
DeepSLM	<b><math>0.9981 \pm 0.0025</math></b>	<b><math>0.3741 \pm 0.2174</math></b>	<b><math>0.7127 \pm 0.1929</math></b>	$0.2279 \pm 0.1061$

private key to decrypt the ciphertext image. This means that the keys generated by the proposed method are unique and our model can resist the attack even if the attacker has all the clues.

### 3.3. Comparison with Existing Encryption Methods

In this section, we benchmark the proposed DeepSLM against a suite of established encryption algorithms, including speckle-based optical cryptosystem,<sup>[25]</sup> Rivest Shamir Adleman (RSA) algorithm,<sup>[54]</sup> and elliptic curve cryptography (ECC)<sup>[55]</sup> algorithm. We assess the quality of the ciphertexts produced by these methods through the following four tests: nonoverlapping template matching, binary matrix rank, Maurer's universal test, and random excursions variant<sup>[56]</sup>: nonoverlapping template matching evaluates if subsequences in the test sequence excessively match nonperiodic templates; binary matrix rank measures the linear independence of fixed-length substrings within the sequence; Maurer's universal test assesses the compressibility of the sequence without information loss, with incompressibility indicating randomness; random excursions variant test examines the deviation in the occurrence frequency of a specific state in a random walk compared to a truly random sequence (if the deviation is large, the sequence is nonrandom).

These tests are quantified by the *P*-value, where a *P*-value is  $\geq 0.01$  signifies a high level of randomness in the ciphertext. According to the results tabulated in Table 8, DeepSLM outperforms the other algorithms on most metrics, with the exception of the random excursion variant test. This indicates that the ciphertexts generated by DeepSLM are characterized by a

superior degree of randomness, underscoring its potential as a robust encryption method.

## 4. Conclusion

In this work, we have proposed a novel image authorization encryption and decryption approach that synergizes optical systems with the prowess of deep learning. The DeepSLM encryption subnetwork is adept at modulating and manipulating optical speckles, which facilitates the creation of ciphertext that embodies various authorization tiers through the integration of conditional data. During the decryption process, these authorization levels precisely dictate the scope of plaintext information accessible to external entities. Our comprehensive experimental evaluations and security analysis affirm that the keys produced by DeepSLM boast an expansive key space, exhibiting exceptional randomness and sensitivity. The resultant ciphertext images are characterized by chaotic pseudorandomness, extraordinary sensitivity to changes, and resilience against a spectrum of attacks. In comparison with conventional encryption algorithms, our method generates ciphertext of superior security. Moreover, the accuracy of the decrypted images is sufficiently high to fulfill the demands of diverse applications, such as face recognition.<sup>[25]</sup> Overall, the fusion of physical optics with artificial intelligence in our system offers the benefits of heightened security, rapid processing, and cost-effectiveness. This study not only proposes but also validates the practicality of permissioned modulation encryption of optical speckles, thus charting a course for the

advancement of learning-enhanced optical speckle encryption in communication, imaging, and data storage domains.

## Acknowledgements

H.H., Q.Z., and H.L. contributed equally to the work. This work was supported by the National Natural Science Foundation of China (NSFC) (grant no. 81930048), Guangdong Science and Technology Commission (grant no. 2019BT02X105), Hong Kong Innovation and Technology Commission (grant nos. GHP/043/19SZ and GHP/044/19GD), Hong Kong Research Grant Council (grant nos. 15217721, 15125724, and C7074-21GF), Shenzhen Science and Technology Innovation Commission (grant no. JCYJ20220818100202005), Hong Kong Polytechnic University (grant nos. P0039517, P0043485, P0045762, and P0048314), and A\*STAR SERC AME program “Nanoantenna Spatial Light Modulators for Next-Generation AR/VR and Holographic Display Technologies”: A18A7b0058.

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

**Haofan Huang:** Conceptualization (lead); Investigation (lead); Methodology (lead); Validation (lead); Writing—original draft (lead). **Qi Zhao:** Data curation (equal); Methodology (equal); Writing—review & editing (equal). **Huanhao Li:** Conceptualization (equal); Methodology (equal); Writing—review & editing (equal). **Yuangong Zheng:** Investigation (equal); Visualization (equal). **Zhipeng Yu:** Investigation (equal); Methodology (supporting). **Tianting Zhong:** Data curation (supporting); Visualization (supporting). **Shengfu Cheng:** Software (supporting). **Chi Man Woo:** Writing—review & editing (supporting). **Yi Gao:** Methodology (supporting); Writing—review & editing (equal). **Honglin Liu:** Methodology (supporting); Writing—review & editing (supporting). **Yuanjin Zheng:** Funding acquisition (supporting); Writing—review & editing (supporting). **Jie Tian:** Methodology (equal); Supervision (equal); Writing—review & editing (equal). **Puxiang Lai:** Funding acquisition (lead); Project administration (lead); Supervision (lead); Writing—review & editing (lead).

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Keywords

authorized encryption and decryption, deep learning, optical speckle, privacy protection, wavefront shaping

Received: February 24, 2024

Revised: August 23, 2024

Published online: September 30, 2024

- [1] SINTEF, I., *Big Data, For Better or Worse: 90% Of World's Data Generated Over Last Two Years*, ScienceDaily, Rockville, MD **2013**.
- [2] P. Vagata, K. Wilfong, *Scaling the Facebook Data Warehouse to 300 PB*, Vol. 10, Facebook Code, Facebook, Menlo Park, CA **2014**.

- [3] M. Lesk, *How Much Information is There in the World?* New Brunswick, NJ **1997**.
- [4] S. Landau, *Science* **2015**, 347, 504.
- [5] T. Huang, H. Xu, H. Wang, H. Huang, Y. Xu, B. Li, S. Hong, G. Feng, S. Kui, G. Liu, D. Jiang, Z.-C. Li, Y. Li, C. Ma, C. Su, W. Wang, R. Li, P. Lai, J. Qiao, *Innovation Med.* **2023**, 1, 1000301.
- [6] H. Berghel, *Computer* **2017**, 50, 72.
- [7] B. Javidi, A. Carnicer, M. Yamaguchi, T. Nomura, E. Pérez-Cabré, M. S. Millán, N. K. Nishchal, R. Torroba, J. F. Barrera, W. He, X. Peng, A. Stern, Y. Rivenson, A. Alfalou, C. Brosseau, C. Guo, J. T. Sheridan, G. Situ, M. Naruse, T. Matsumoto, I. Juvells, E. Tajahuerce, J. Lancis, W. Chen, X. Chen, P. W. H. Pinkse, A. P. Mosk, A. Markman, *J. Opt.* **2016**, 18, 083001.
- [8] P. Zheng, Q. Dai, Z. Li, Z. Ye, J. Xiong, H.-C. Liu, G. Zheng, S. Zhang, *Sci. Adv.* **2021**, 7, eabg0363.
- [9] O. Matoba, T. Nomura, E. Perez-Cabre, M. I. S. Millan, B. Javidi, *Proceedings of the IEEE* **2009**, 97, 1128.
- [10] Z. Yu, T. Zhong, H. Li, H. Li, C. Man Woo, S. Cheng, S. Jiao, H. Liu, C. Lu, P. Lai, *Photonics Res.* **2024**, 12, 587.
- [11] H. Li, Z. Yu, Q. Zhao, T. Zhong, P. Lai, *Innovation* **2022**, 3, 100252.
- [12] Z. Yu, H. Li, W. Zhao, P.-S. Huang, Y.-T. Lin, J. Yao, W. Li, Q. Zhao, P. C. Wu, B. Li, P. Genevet, Q. Song, P. Lai, *Nat. Commun.* **2024**, 15, 2607.
- [13] P. Refregier, B. Javidi, *Opt. Lett.* **1995**, 20, 767.
- [14] X. He, H. Tao, L. Zhang, X. Yuan, C. Liu, J. Zhu, *IEEE Photonics J.* **2019**, 11, 1.
- [15] B. Javidi, T. Nomura, *Opt. Lett.* **2000**, 25, 28.
- [16] B. Hennelly, J. T. Sheridan, *Opt. Commun.* **2003**, 226, 61.
- [17] Y. Zhang, L. Y. Zhang, J. Zhou, L. Liu, F. Chen, X. He, *IEEE Access* **2016**, 4, 2507.
- [18] N.-R. Zhou, L.-X. Huang, L.-H. Gong, Q.-W. Zeng, *Quantum Inf. Process.* **2020**, 19, 284.
- [19] L.-H. Gong, H.-X. Luo, R.-Q. Wu, N.-R. Zhou, *Phys. A* **2022**, 591, 126793.
- [20] L. Zhou, Y. Xiao, W. Chen, *Opt. Lett.* **2020**, 45, 5279.
- [21] X. Wang, W. Wang, H. Wei, B. Xu, C. Dai, *Opt. Lett.* **2021**, 46, 5794.
- [22] X. Wang, H. Wei, M. Jin, B. Xu, J. Chen, *Opt. Express* **2022**, 30, 11165.
- [23] L. Zhou, Y. Xiao, W. Chen, *Opt. Lasers Eng.* **2021**, 141, 106570.
- [24] F. Feng, J. Hu, Z. Guo, J.-A. Gan, P.-F. Chen, G. Chen, C. Min, X. Yuan, M. Somekh, *ACS Photonics* **2022**, 9, 820.
- [25] Q. Zhao, H. Li, Z. Yu, C. M. Woo, T. Zhong, S. Cheng, Y. Zheng, H. Liu, J. Tian, P. Lai, *Adv. Sci.* **2022**, 9, 2202407.
- [26] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*, Roberts and Company Publishers, Greenwood Village, CO **2007**.
- [27] V. Himthani, V. S. Dhaka, M. Kaur, D. Singh, H.-N. Lee, *IEEE Access* **2022**, 10, 98360.
- [28] C. Li, F. Zhao, C. Liu, L. Lei, J. Zhang, *Security and Communication Networks* **2019**, 2019, 8132547.
- [29] M. Kaur, S. Singh, M. Kaur, *Math. Probl. Eng.* **2021**, 2021, 5012496.
- [30] H. Li, Z. Yu, Q. Zhao, Y. Luo, S. Cheng, T. Zhong, C. M. Woo, H. Liu, L. V. Wang, Y. Zheng, P. Lai, *Photonics Res.* **2023**, 11, 631.
- [31] E. Akkermans, G. Montambaux, *Mesoscopic Physics of Electrons and Photons*, Cambridge University Press, Cambridge, England **2007**.
- [32] Co-operation, O.f.E. and Development, *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, Paris, France **1980**.
- [33] Z. Yu, H. Li, T. Zhong, J.-H. Park, S. Cheng, C. M. Woo, Q. Zhao, J. Yao, Y. Zhou, X. Huang, W. Pang, H. Yoon, Y. Shen, H. Liu, Y. Zheng, Y. K. Park, L. V. Wang, P. Lai, *Innovation* **2022**, 3.
- [34] X. Zhang, J. Gao, Y. Gan, C. Song, D. Zhang, S. Zhuang, S. Han, P. Lai, H. Liu, *Photonix* **2023**, 4, 10.
- [35] G. Qu, W. Yang, Q. Song, Y. Liu, C.-W. Qiu, J. Han, D.-P. Tsai, S. Xiao, *Nat. Commun.* **2020**, 11, 5484.



- [36] M. Lyu, H. Wang, G. Li, S. Zheng, G. Situ, *Adv. Photonics* **2019**, 1, 036002.
- [37] O. Katz, P. Heidmann, M. Fink, S. Gigan, *Nat. Photonics* **2014**, 8, 784.
- [38] T. Karras, S. Laine, T. Aila, in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2019**.
- [39] Flickr-Faces-HQ Dataset, <https://github.com/NVlabs/ffhq-dataset> (accessed: September 2024).
- [40] J. Wang, et al., *Int. J. Comput. Vis.* **2024**, 1.
- [41] S. Ioffe, C. Szegedy, in *Int. Conf. Machine Learning*, JMLR.org, Lille, France **2015**.
- [42] H. De Vries, F. Strub, J. Mary, H. Larochelle, O. Pietquin, A. C. Courville, *Adv. Neural Inf. Process. Syst.* **2017**, 30.
- [43] Y. Ding, F. Tan, Z. Qin, M. Cao, K. R. Choo, Z. Qin, *IEEE Trans. Neural Networks Learn. Syst.* **2021**, 33, 4915.
- [44] Y. Taigman, A. Polyak, L. Wolf, in *Int. Conf. on Learning Representations* **2022**.
- [45] J. H. Lim, J. C. Ye (Preprint), arXiv:1705.02894, v1, Submitted: May **2017**.
- [46] T. Miyato, M. Koyama, in *Int. Conf. on Learning Representations* **2018**.
- [47] O. Ronneberger, P. Fischer, T. Brox, in *Int. Conf. Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, Switzerland **2015**.
- [48] S. Li, M. Deng, J. Lee, A. Sinha, G. Barbastathis, *Optica* **2018**, 5, 803.
- [49] D. P. Kingma, J. Ba (Preprint), arXiv:1412.6980, v1, Submitted: Dec. **2014**.
- [50] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, *Adv. Neural Inf. Process. Syst.* **2019**, 32.
- [51] A. A. Tamimi, A. M. Abdalla, M. M. Abdallah, in *Advances in Computer Vision and Computational Biology: Proc. from IPCV'20, HIMS'20, BIOCOMP'20, and BIOENG'20*, Springer, Cham, Switzerland **2021**, pp. 271–278.
- [52] B. Schneier, *Applied Cryptography*, John Wiley & Sons Inc., New York, NY **1996**.
- [53] Y. Wu, J. P. Noonan, S. Agaian, *J. Sel. Areas Telecommun.* **2011**, 1, 31.
- [54] Y. Desmedt, A. M. Odlyzko, in *Conf. Theory and Application of Cryptographic Techniques*, Springer, Berlin, Heidelberg **1985**.
- [55] J. W. Bos, J. A. Halderman, N. Heninger, J. Moore, M. Naehrig, E. Wustrow, in *Int. Conf. Financial Cryptography and Data Security*, Springer, Berlin, Heidelberg **2014**.
- [56] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, S. Vo, *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*, Vol. 22, US Department of Commerce, Technology Administration, National Institute of Standards and Technology, Gaithersburg, MD **2001**.