ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20240801001893
Original scientific paper

GAN-Based Facial Information Protection for IoT Using Transfer Probability Models

Ye QIN, Yaowu KANG*

Abstract: Facial information in the Internet of Things (IoT) faces great security challenges. This study proposes a novel facial information protection model combining Deep Convolutional Generative Adversarial Networks (DCGAN) with Transfer Probability Models (TPM). The model generates high-quality virtual face images while preserving key features of the original image. The results demonstrated that the model performed well in terms of encryption-decryption error (0.05), speed (632.5 Mbit/s for encryption and 583.5 Mbit/s for decryption), resource consumption (21.1%), and latency (11.5%) compared with existing methods. The model achieved more than 90% privacy protection for identity, facial expression, shape, and gesture, proving its effectiveness in facial information protection for IoT.

Keywords: deep convolutional generative adversarial network; facial features; information; protection transition probability model

1 INTRODUCTION

Traditional personal information authentication relies on character passwords and tokens, which are easily forgotten and stolen. The use of biometrics for the authentication of personal information is therefore becoming increasingly prevalent among the general public. Internet of Things (IoT) devices offer convenience and the potential risk of privacy leakage due to their capacity for continuous data collection and processing of large amounts of personal data. Among these personal data, facial information is susceptible because it is not only used for authentication but also often associated with other private information. This makes protecting facial information a non-negligible part of IoT security [1, 2]. However, as technology advances, facial authentication has started to reveal certain shortcomings, such as vulnerabilities to facial information forgery and leakage. The main Facial Information Protection (FIP) techniques at this stage include homomorphic encryption, differential privacy, and federated learning. Homomorphic encryption is an encryption method that allows operations to be performed directly on the encrypted data, ensuring that the data can still be computed in an encrypted state. Differential privacy technology protects the privacy of individual data by adding random noise to the dataset, while still allowing analysis of the overall data. While these techniques improve the security of facial information to some extent, they still have significant shortcomings. For example, although homomorphic encryption allows operations on encrypted data, it has high computational complexity and slow processing speed, making it unsuitable for large-scale real-time applications. Differential privacy protects privacy by adding noise, but it may lead to data distortion and affect recognition accuracy. In addition, existing methods often fail to balance image quality and information security when facing highly complex image feature retention and privacy protection, resulting in poor image restoration and easy privacy leakage [3]. Generative Adversarial Networks (GANs) have gained widespread use in FIP due to their superior performance in image generation and data enhancement [4]. Specifically, the Deep Convolutional GAN (DCGAN) is known for generating high-quality virtual face images through adversarial training between generators and discriminators [5]. However, DCGAN still has some limitations in privacy protection. For instance, the feature information of the generated image may potentially reveal a degree of the original image's privacy. For this reason, the study innovatively introduces deep Residual Network (ResNet) for convolutional improvement and adds Transfer Probability Model (TPM) to enhance the image feature retention. This method aims to fill the gap in the balance between processing speed, feature preservation, and privacy protection in existing technologies by generating high-quality virtual face images and ensuring the preservation of key features of the original image after decryption. Based on this, more efficient and secure solutions can be provided to address the increasingly severe facial information security challenges in the IoT environment.

2 LITERATURE REVIEW

Facial information, as a kind of unique and unchangeable personal identification information, may lead to serious privacy leakage and security risks once it is acquired and utilized by unscrupulous elements. Xie Y. et al. In order to get rid of the storage privacy nuisance of local or mobile terminals for face information recognition systems, the research team proposed a generalized privacy-preserving framework for edge-based face recognition systems. Experimental results show that the framework effectively achieves the best balance between usability and privacy protection for face information recognition and transmission [6]. However, although the method performed well in video surveillance, it still lacked the ability to process static images, and the blurring process might affect the image recognition accuracy in practical applications. Sardar A. et al. identified significant vulnerabilities in biometric recognition systems that utilize facial information. In response, they proposed a facial biometric template protection model that integrates the concept of revocable biometric features. This model enhanced facial protection through the use of feature vectors, proving to be more effective than traditional methods [7]. Nevertheless, the model still relied on biological templates, and once the templates were breached, the security would be greatly threatened. In addition, there were limitations in the selection and combination of feature vectors. Dhinakaran D. et al. found that the protection efficiency of personal information,

especially facial data, within the IoT is extremely poor when relying on cryptographic tokens alone. To address this, they combined the whale optimization algorithm with frequent term mining to develop a new FIP system. This system effectively safeguarded confidential information disclosed by individual users, offering high-quality protection [8]. The system effectively improved the quality of protection of user disclosures, but its complex optimization process might increase the computational overhead, limiting the feasibility of large-scale applications. Angel D. et al. proposed a model for protecting personal medical information using Bloom topology and a linear discriminant analysis algorithm. This model enhanced the security of personal electronic health records within medical information systems and could effectively counter common proactive network attacks, with a data protection effectiveness of up to 94.8% [9]. Although the method performed well in medical information protection, it was not originally designed to address the unique challenges of facial information and may suffer from a lack of applicability when applied directly to FIP.

GANs are a cutting-edge deep learning techniques capable of generating high-quality virtual data through the adversarial training of two networks [10]. In recent years, GANs have achieved remarkable success in areas such as image generation and data augmentation. For instance, You, A. et al. combined fuzzy perception and GANs and tried to apply them to computer image classification and finally proposed a novel classification method. Experimental results show that the method is more effective and robust in video surveillance image classification [11]. However, GANs still faced the risk of training instability and feature leakage when generating high-quality virtual images, which also limits their application in the broader field of FIP. Wang Y. et al. identified the challenge of training image deblurring networks using real-world paired blurry or clean images, which are difficult to capture. To address this, they proposed a novel image deblurring patch GAN model that combines patch estimation with GAN. This method provided high robustness in fuzzy image restoration and preserved more comprehensive details, demonstrating the broad applicability of GANs in FIP [12]. Although the method achieved good results in image deblurring, the deblurring process might inadvertently recover or expose sensitive information when protecting privacy. This raised a potential conflict between information protection and image recovery. Mahmoudinejad et al. noted that while style-based GAN methods can generate highly realistic facial images, controlling the features of the generated faces in a meaningful and clear way often proved difficult. To enhance the accuracy of facial image processing, they developed a disentanglement GAN model. This model demonstrated excellent realism and nonlinear 3D deformability in generating facial information [13]. Although the model performed well in facial information generation, its effectiveness in information protection was limited, especially in generating images that made it difficult to avoid partial information leakage from the original image. Chen B. et al. observed that the generalization ability of existing FIP techniques using GANs needed improvement. To address this, they enhanced GAN by integrating ArcFace loss and proposed a new method for protecting facial information. The highest testing effectiveness of this method on natural datasets reached 94.2% [14]. Although it performed well in tests on natural datasets, the model's ability to generalize to more complex and diverse datasets required further validation.

In summary, existing technologies have achieved notable progress in FIP, enhancing the efficiency of blurring protection and privacy preservation for facial images. However, limitations remain, including challenges in facial feature control, processing accuracy, and generalization ability. To address these challenges, the study proposes a novel FIP method combining improved DCGAN and TPM. This model not only effectively protects the privacy of facial information but also achieves a good balance between image quality and feature retention. The study provides a new solution to address the information security challenges in IoT environments.

3 RESEARCH METHODOLOGY

To address the issue of facial information security in the protection of personal information in the IoT, this study first introduces DCGAN as the basic model. However, this model has limitations when dealing with complex features. To address the issues of gradient vanishing and feature preservation, ResNet is introduced and improved the efficiency of network training and image feature preservation through skip connections. In addition, to ensure that the decrypted face image can accurately retain identity and attribute features, the study combines the TPM. This model further optimizes the decryption effect through the calculation of feature transfer probability. Ultimately, a novel FIP model combining DCGAN, ResNet, and TPM is proposed, which significantly improves the quality and security of image generation. The overall flow of the method is shown in Fig. 1.

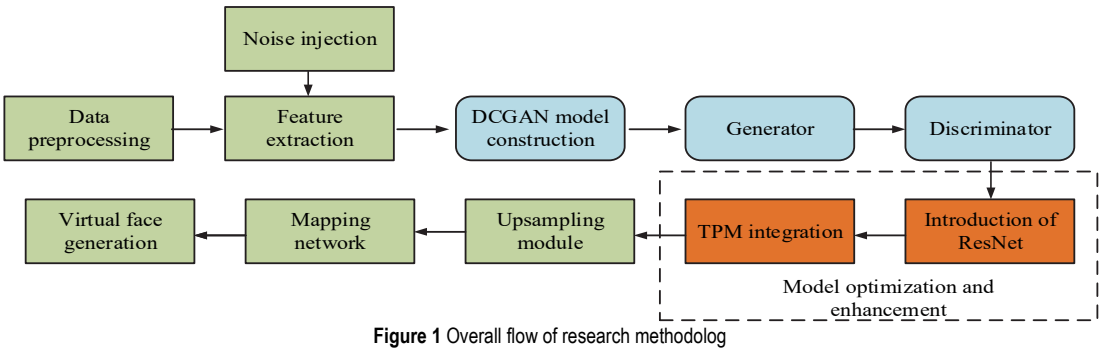

Figure 1 Overall flow of research methodolog

3.1 Construction of FIP Model Based on GAN

GANs comprise two adversarial networks: a generator and a discriminator. The generator's role is to produce realistic virtual data, while the discriminator's role is to distinguish between real data and the data generated by the generator [15, 16]. Through continuous adversarial training, the generator improves its ability to create data that becomes increasingly indistinguishable from real data. The calculation formula for the loss function in the generator is represented in Eq. (1).

$$L_G = -\log\left(D\left(G(x)\right)\right) \tag{1}$$

In Eq. (1), $G(x)$ represents the face image processed by the generator, where x is the original facial image. $D(x)$ is the evaluation result of the discriminator on the original image. The term L_G refers to the objective maximization of the generator. The corresponding loss function for the discriminator, which aims to correctly distinguish between real and generated images, is provided in Eq. (2).

$$L_D = -\left[\log\left(D(x)\right) + \log\left(1 - D\left(G(x)\right)\right)\right] \tag{2}$$

In Eq. (2), $D\left(G(x)\right)$ represents the evaluation result of the discriminator on the image generated by the generator. Although traditional GANs increase the flexibility of the model, they also introduce training instability, sometimes even leading to the collapse of the model training process [17, 18]. Therefore, this study introduces DCGAN. Compared with other GAN architectures, DCGAN has better stability and generation quality in processing image generation tasks. Specifically, DCGAN uses convolutional and anti-convolutional layers instead of the fully connected layers in traditional GANs, which is more suitable for capturing the local structure and spatial information of images. This allows it to generate high-quality images while maintaining high training stability. The DCGAN structure is shown in Fig. 2.

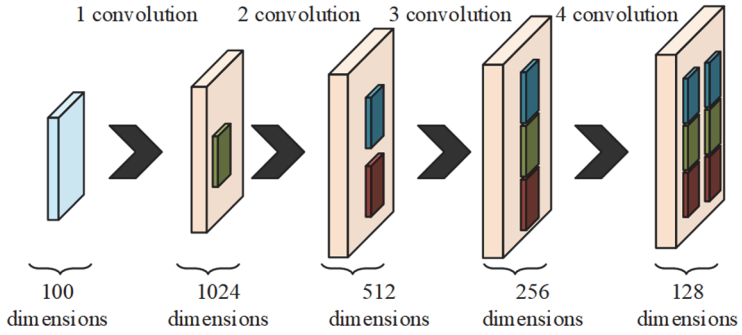

Figure 2 DCGAN structure diagram

In Fig. 2, the DCGAN generator is mainly divided into four convolutional layers. The first layer expands the 100 dimensional noise to 1024 dimensions to capture preliminary features. The second layer shrinks to 512 dimensions to extract high-level features. The third layer shrinks to 256 dimensions to enhance the image structure. The fourth layer ultimately outputs a 128 dimensional high-resolution image. Each layer of convolution is followed by a combination of batch normalization and activation functions to ensure generation quality and training stability. The function of the generator is represented by Eq. (3).

$$L_G` = -E_{z\sim pz(z)}\left[\log\left(D`\left(G`(z)\right)\right)\right] \tag{3}$$

In Eq. (3), E represents the expectation operator, typically used to calculate the average or expected value. x denotes a data sample from a real dataset, while z is a noise sample drawn from the latent spatial prior distribution. $G`(z)$ refers to the data generated by the generator using the noise sample $z$. $D`(x)$ is the output of the discriminator when evaluating the real sample x, which reflects the probability that the data is real. $D`\left(G`(z)\right)$ is the output of the discriminator when evaluating the data generated by the generator. The function of the discriminator is then represented by Eq. (4).

$$\begin{aligned} L_D` = &-E_{x\sim pdata(x)}\left[\log(D`(x)\right] - \\ &-E_{z\sim pz(z)}\left[1 - D\left(G`(x)\right)\right] \end{aligned} \tag{4}$$

Compared to traditional GANs, DCGAN utilizes convolutional and deconvolutional layers in place of fully connected layers, enabling it to better capture the local structure and spatial information of images. Additionally, the performance of DCGAN models tends to improve as the depth of network training increases. However, when the network's hierarchy surpasses a certain threshold, the model's performance may plateau or even degrade. To address this issue, this study introduces the ResNet

architecture. The ResNet enhances traditional convolutional networks by incorporating skip connections, which directly connect the input and output of a layer. This approach helps mitigate issues such as vanishing gradients and allows for deeper network architectures without a corresponding decrease in performance [19, 20]. The core formula governing ResNet's operations is provided in Eq. (5).

$$y = F\left(x,\left(W_i\right)\right)+W_s x \tag{5}$$

In Eq. (5), y represents the output of the residual block. F is the residual function. W_s is the transformation matrix, and W_i denotes the weights of each layer within the residual unit. In summary, this study proposes a novel FIP model that combines the strengths of DCGAN and ResNet, leading to an improved DCGAN architecture. The structure of this model is depicted in Fig. 3.

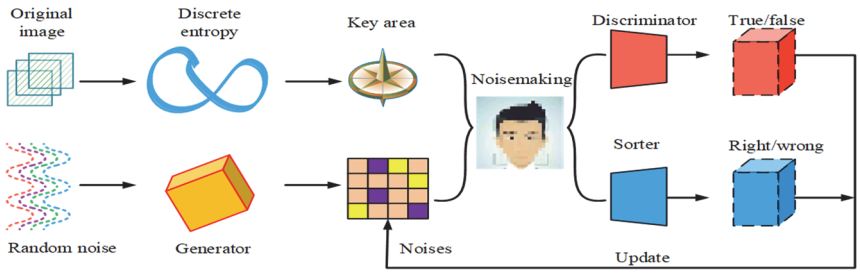

Figure 3 The improved DCGAN model structure for FIP

In Fig. 3, the process begins by extracting the key facial feature regions from the original face image using discrete entropy calculations. Random noise is then added to the original face image. Next, the denoised facial image, along with the extracted key facial regions, is input into the generator, which generates facial denoised features. These features, along with the original image, are then input into the discriminator, which distinguishes between true (original) and false (noisy) images. The goal is for the discriminator to accurately identify the original image as true and the noisy image as false. The detailed structures of the generator and discriminator are illustrated in Fig. 4.

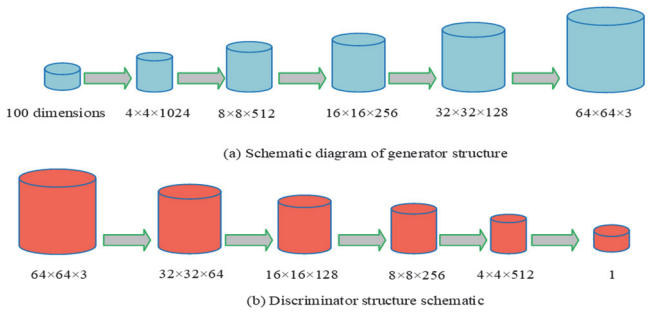

Figure 4 The generator and discriminator structure of the improved DCGAN

Fig. 4a and Fig. 4b depict the structural diagrams of the generator and discriminator, respectively. The generator begins by expanding and reshaping a 100 dimensional noise vector into a 4 × 4 × 1024 feature map. This map then undergoes four layers of convolution and transpose convolution operations, progressively generating images of sizes 8 × 8 × 512, 16 × 16 × 256, and 32 × 32 × 128, ultimately producing a final output of 64 × 64 × 3. The discriminator starts with an input image of 64 × 64 × 3 and progressively reduces the feature map size through four convolutional layers, resulting in feature maps of 32 × 32 × 64, 16 × 16 × 128, 8 × 8 × 256, and 4 × 4 × 512. The process culminates in a scalar output via a fully connected layer, which determines the authenticity of the image. Compared to traditional DCGAN, this model's generator produces higher-resolution images by increasing the number of layers and incorporating transpose convolution operations. The discriminator improves its ability to distinguish between real and generated images through a series of layered convolutions and a fully connected final layer.

3.2 Strengthening GAN FIP in Conjunction with TPM

Although enhancing DCGAN can provide denoising protection for facial image information, decrypted facial images over extended periods tend to lose the correlation between the original image and its attribute features. This results in decrypted images that do not fully retain facial identity information and attribute characteristics [21, 22]. To address this issue, this study introduces the TPM to facilitate the transfer and fusion of attribute and identity features. Compared with other methods, TPM can effectively retain the key features and attributes of the face image during the decryption process. TPM can ensure the accuracy and integrity of the decrypted image by calculating the transfer probability between the identity features and the attribute features, as well as improve the effect of privacy protection [23, 24]. The calculation of feature transfer within TPM is expressed in Eq. (6).

$$P\left(T_i \middle| T_{i-1}, T_{i-2}, \ldots, T_{i-n}\right) = \frac{P\left(T_i, T_{i-1}, T_{i-2}, \ldots, T_{i-n}\right)}{P\left(T_{i-1}, T_{i-2}, \ldots, T_{i-n}\right)} \quad (6)$$

In Eq. (6), T denotes the feature. $P\left(T_i \middle| T_{i-1}, T_{i-2}, \ldots, T_{i-n}\right)$ is the probability of feature T appearing given the first n features $T_{i-1}, T_{i-2}, \ldots, T_{i-n}$. The formula for attribute fusion probability is Eq. (7).

$$P\left(A_i \middle| I_i\right) = P\left(I_i \middle| A_i\right) \cdot \frac{P\left(A_i\right)}{P\left(I_i\right)} \quad (7)$$

In Eq. (7), $P\left(A_i \middle| I_i\right)$ and $P\left(I_i \middle| A_i\right)$ are the probabilities of attribute feature A_i and identity feature I_i appearing under the conditions of identity features I_i and A_i. $P\left(A_i\right)$ and $P\left(I_i\right)$ are the marginal probabilities of attribute features and identity features. The expression of fused features is Eq. (8).

$$P\left(T_i, A_i\right) = P\left(T_i \middle| A_i\right) \cdot P\left(A_i\right) \quad (8)$$

By integrating Eq. (6), Eq. (7), and Eq. (8), the encoder and decoder for multi-scale feature and attribute transfer are illustrated in Fig. 5.

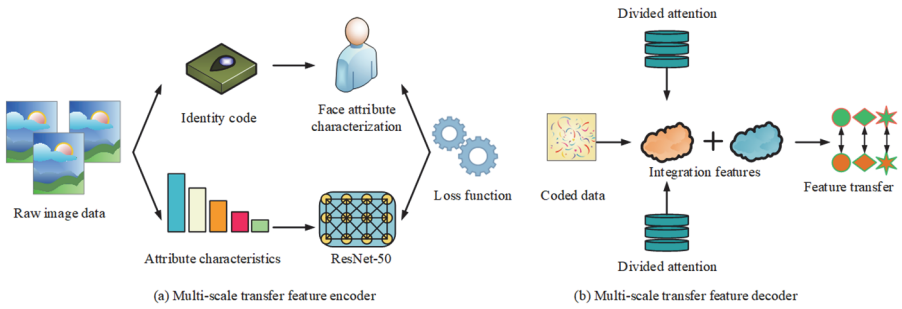

Figure 5 Schematic of encoder and decoder for multi-scale transfer features

Fig. 5a and Fig. 5b depict the multi-scale transfer feature encoder and decoder. ResNet-50 serves as the backbone network for the feature encoder, which encodes facial attribute features and extracts facial identity features through probability calculations and identity mapping. The decoder, upon receiving the identity and attribute features encoded by the encoder, transfers these features through the fusion process in the cross-attention module. However, due to the multidimensional nature and complex origins of the features, the transfer of fused features often results in the appearance of small artifacts or spots [25, 26]. To address this issue, the study improves the upsampling and mapping network of the model. The schematic diagram of these improvements is shown in Fig. 6.

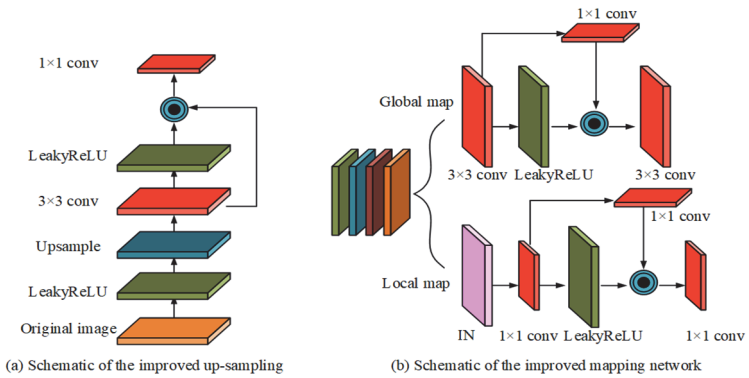

Figure 6 Schematic diagram of the new upsampling and mapping network structure

In Fig. 6, the improved upsampling module integrates the LeakyReLU activation function with convolutional layers, employing upsampling and 3 × 3 convolution operations to produce high-resolution feature maps. In the mapping network, the global mapping component uses a structure that combines two consecutive 3 × 3 convolutions with Instance Normalization (IN). The local mapping component further adjusts and enhances features through a combination of 1 × 1 convolution and IN. In summary, this study combines the improved DCGAN with feature transition probability to propose a novel IoT facial

information security protection model. The workflow of this model is illustrated in Fig. 7.

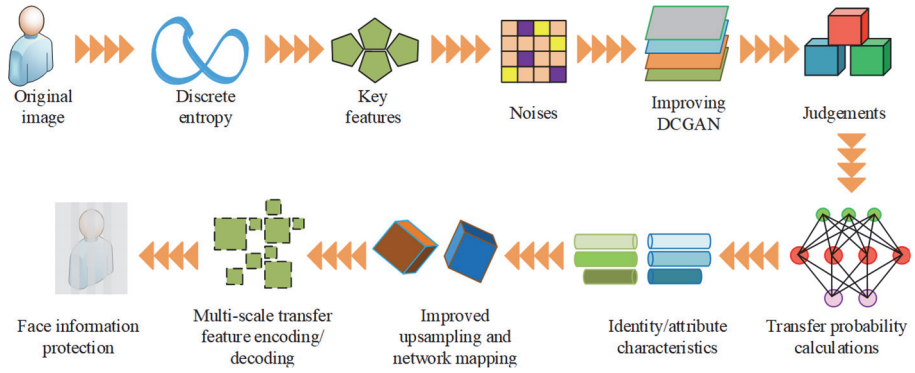

Figure 7 A novel facial information protection model for the IoT

In Fig. 7, the process begins with the extraction of key feature regions from the original face image using discrete entropy calculation. In the second step, random noise is added to the original face image as an initial layer of protection. In the preprocessing step, images are cropped and scaled to unify the resolution, and histogram equalization is applied to normalize image brightness and contrast. In addition, to enhance data diversity, data enhancement techniques including random flipping, rotation, and noise injection are applied to simulate diverse image conditions in real scenes. The third step involves inputting the denoised image into an improved DCGAN generator to produce a face image with noisy features. In the fourth step, the generated noisy image is processed using the TPM to ensure that the decrypted face image retains the key features and attributes of the original image. This step also includes enhancing the resolution and details of the generated image through an improved upsampling and mapping network. In the fifth step, the processed image is input into a multi-scale transfer feature encoder to extract multi-scale features. The sixth and final steps involve fusing these multi-scale features through a decoder, resulting in the generation of a high-quality facial image. During the model training process, the learning rate of the generator and discriminator is set to 0.0002. To prevent overfitting, the batch size is set to 128 and the Adam optimizer is used for parameter updating. The latent space dimension is chosen to be 100. In addition, the number of residual blocks is set to be 5 when the ResNet is introduced. For TPM, the weighting factor of its transfer probability is set to be 0.5. The improved generator function used in this process is described by Eq. (9).

$$L_G^* = E_{mask_{init} \sim Pmask_{init}(mask_{init})}\left[\log\left(1 - D\left(e_{G(mask_{init})}\right)\right)\right] \quad (9)$$

In Eq. (9), L_G^* represents the improved generator loss function, and E denotes the expectation operator. The terms $mask_{init}$ and $P_{mask_{init}}$ refer to the initial masks and their corresponding probability distributions, respectively. $e_{G(mask_{init})}$ is the output generated by the generator G for the initial mask $mask_{init}$. The symbol D represents the discriminator. The improved discriminator function is provided in Eq. (10).

$$L_D^* = E_{e \sim P_{data}}(e)[\log(D(e))] + \\ + E_{mask_{init} \sim P_{mask_{init}}}(mask_{init})\left[\log(1 - D(e_{G(mask_{init})}))\right] \quad (10)$$

In Eq. (10), e represents the original facial image, and P_{data} is the probability distribution of e. At this stage, the generator and discriminator engage in an adversarial process to calculate the Euclidean distance between them. A larger distance suggests that the performance of the classifier is better. The calculation of the face classifier is described in Eq. (11).

$$L_{target} = -\log[ED(e, e_{G(mask_{init})})] \quad (11)$$

In Eq. (11), $ED(\_)$ is the Euclidean distance between the two.

4 Results and Discussion

This study first established an appropriate experimental environment to evaluate the performance of the proposed model through ablation testing. Additionally, the model was tested for encryption and decryption errors to assess its reliability. Performance indicators such as running speed, processing time, and resource consumption rate were used to further evaluate the model. Simulation tests were also conducted using real facial images, with histograms generated for each model to visualize the results. Finally, multi-criteria testing was performed to verify the authenticity and effectiveness of the model.

4.1 Performance Testing of IoT-FIP Model

The experimental environment for this study was set up using an Intel Core i7-9700 CPU @ 3.00 GHz × 32 and an NVIDIA GeForce RTX 3060 GPU. All data processing was implemented in MATLAB, with Python 3.7 as the programming language. The Chinese Academy of Sciences WebFace Database (CASIA-WebFace) and the Celebrity Face Attributes Data Set (CelebA) were used as test data sources. CASIA-WebFace was primarily used for large-scale facial recognition research, featuring images captured from IoT devices that include extensive identity information across various scenarios. CelebA, on the other hand, comprised images sourced from celebrity photos, covering a wide range of facial attributes. These two types of data were chosen because the CASIA-WebFace dataset is rich in identity labels and can provide diversity samples to help the model improve its generalization ability when recognizing and validating different facial features. The CelebA dataset covers a wide range of facial attributes such as gender, age, hairstyle, and expression, making it well suited for studying tasks related to facial features and attributes. The study first evaluated the new IoT-FIP model using ablation testing. The results are presented in Fig. 8.

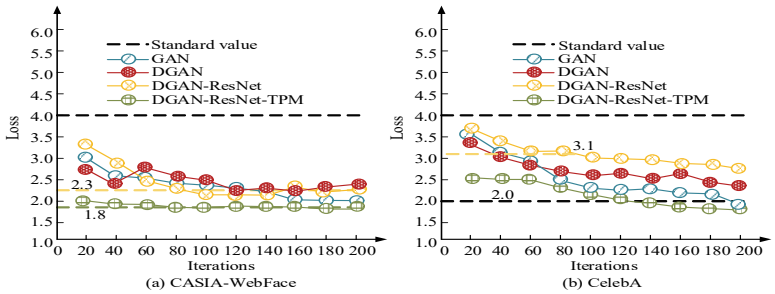

Figure 8 Novel FIP model ablation test

The ablation test results shown in Fig. 8a and Fig. 8b indicate that the difference in loss value between DCGAN and GAN is not significant. However, there is a substantial performance improvement when ResNet is introduced. This enhancement becomes even more pronounced with the addition of the TPM, resulting in an average loss of 1.8 on the CASIA-WebFace dataset and 2.0 on the CelebA dataset. These findings demonstrate that the integration of various mechanisms and new modules positively impacts the overall model performance. However, although these improvements enhance the performance of the model, the introduction of more layers of network structure may increase the computational complexity and training time, which needs to be weighed in practical applications. Furthermore, future research can further optimize these mechanisms to reduce computational overhead and improve the efficiency of the model. Given this, the study compares the error rates after image encryption and decryption with more advanced models of the same type as DCGAN, such as Progressive Growth of GAN (PGGAN), Boundary Equilibrium GAN (BEGAN), and Self-Attention GAN (SAGAN). The specific results of this comparison are presented in Fig. 9.

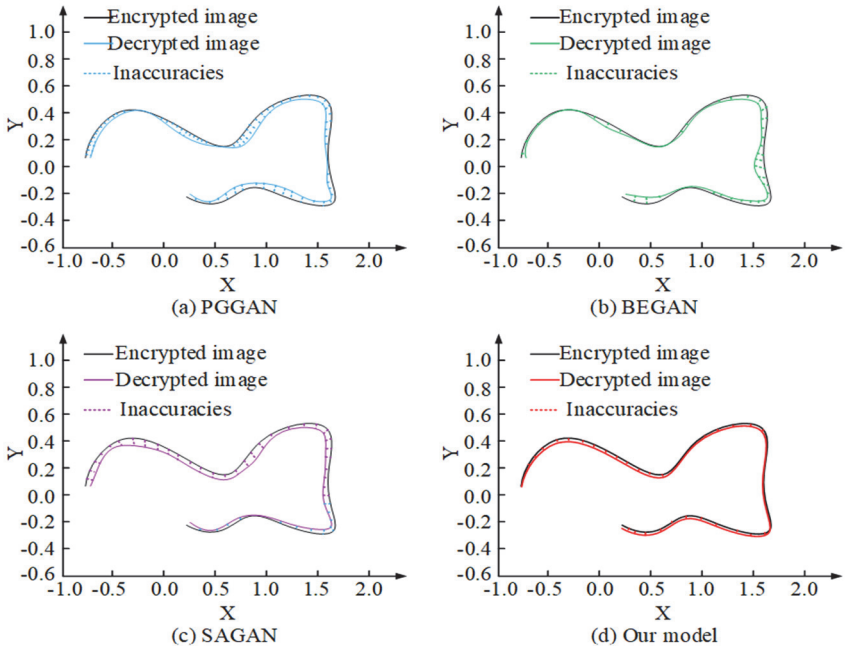

Figure 9 Encryption-decryption reduction test of face images with different models

Fig. 9a to Fig. 9d present the image restoration test results for PGGAN, BEGAN, SAGAN, and the proposed research model. For the same facial image, the overall degree of image restoration after encryption and decryption is similar across the four models. However, in terms of finer details, the research model achieves a higher degree of restoration, with a minimum error of 0.05 in the restored images. In comparison, the minimum image restoration errors for the PGGAN, BEGAN, and SAGAN models are 0.1, 0.1, and 0.08, respectively. This suggests that the structural improvements in the upsampling and mapping network of DCGAN have enhanced the retention of facial attribute features and identity features during encryption. However, it is also important to consider that this increase in accuracy may come at the cost of higher computation and storage costs, which may become a bottleneck in large-scale applications. Therefore, techniques that reduce cost while maintaining high reduction are directions that should continue to be explored in subsequent research. The study continues to test the encryption speed, decryption speed, resource consumption rate, and latency rate as metrics. The results are statistically analyzed using SPSS17.0 software and the results of the tests are shown in Tab. 1.

Table 1 Comparison results of FIP efficiency of different models

Data set	Model	Encryption speed / Mbit/s	Decryption speed / Mbit/s	Resource consumption rate / %	Delay rate / %	*P*
CASIA-WebFace	PGGAN	458.2	412.1	37.4	27.4	0.003
	BEGAN	527.6	587.6	31.6	17.6	0.004
	SAGAN	441.7	428.2	45.7	13.4	0.003
	Research model	632.5	583.5	21.4	11.5	0.006
CelebA	PGGAN	553.7	455.4	33.4	24.5	0.002
	BEGAN	507.2	478.6	29.8	23.6	0.006
	SAGAN	458.9	431.2	35.6	19.8	0.003
	Research model	607.3	497.5	21.1	15.3	0.002

In Tab. 1, the new model achieves encryption and decryption speeds of 632.5 Mbit/s and 583.5 Mbit/s on the CASIA-WebFace dataset, and 607.3 Mbit/s and 497.5 Mbit/s on the CelebA dataset, which are significantly higher than those of the other models. This shows that the new model possesses higher efficiency in processing large-scale data and can meet the demand for real-time and high throughput in IoT environments. Regarding resource consumption, the model demonstrates the lowest rates on both datasets, at 21.4% and 21.1%, respectively. It also records the lowest latency rates of 11.5% and 15.3%, showing its efficient resource utilization and exceptionally low latency. Through statistical analysis, the *P*-values of the new model on all metrics are lower than 0.05 ($P < 0.05$), indicating that these performance enhancements are statistically significant. These results indicate that combining the improved DCGAN with the transmission probability model achieves the research goal of improving privacy protection while maintaining high image quality. Moreover, this approach successfully reduces computational overhead and latency, enhancing the practicality of the model. This has significant practical implications for IoT security. This means that stronger privacy protection can be provided without sacrificing performance, making IoT devices more secure and adaptable to the needs of large-scale application environments.

4.2 Simulation Testing of IoT-FIP Model

To verify the actual performance of the IoT-FIP model, real face images from the CASIA-WebFace and CelebA datasets were selected for testing. Advanced methods for face image encryption and protection, such as Homomorphic Encryption (HE), Differential Privacy (DP), and Secure Multi-Party Computation (SMPC), were also introduced for comparison. The results of these comparisons are shown in Fig. 10.

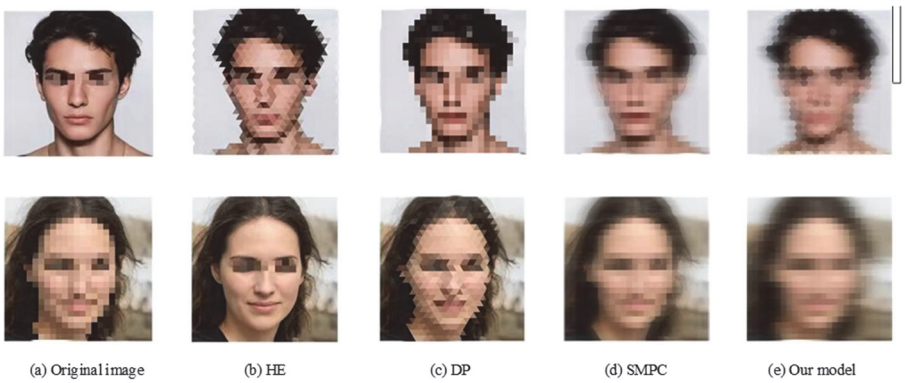

Figure 10 Comparison of encryption effects of different image encryption models

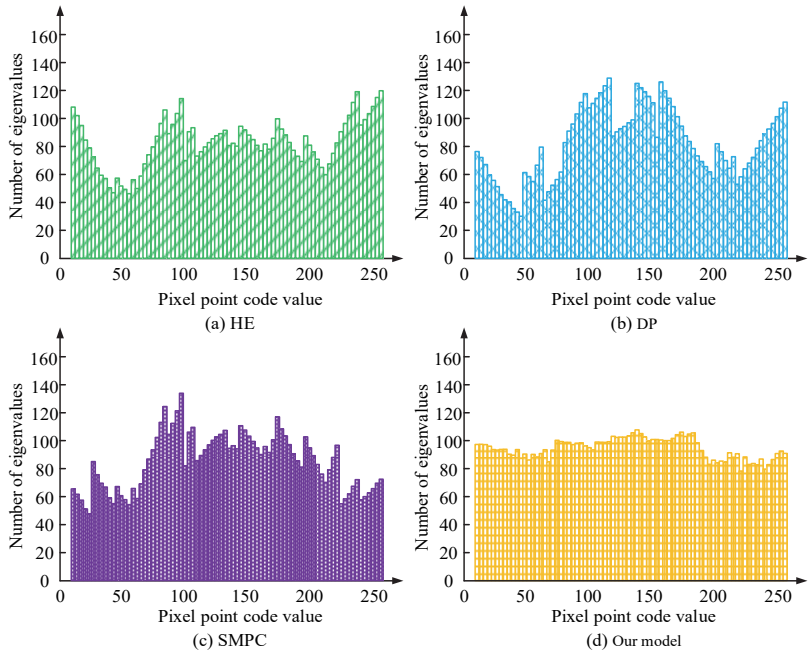

Figure 11 Histogram test results for different models

Fig. 10a presents two original facial photos, while Fig. 10b to Fig. 10e display the image encryption effects using HE, DP, SMPC, and the research model, respectively. Although HE and DP provide privacy protection, they still retain some facial features, which poses a risk of recognition. SMPC offers strong privacy protection but results in high image blurriness, which can negatively impact image usability. In contrast, the research model strikes a better balance between privacy protection and image quality. The encrypted images produced by this model maintain clearer contours and details while effectively blurring key features, thereby enhancing privacy protection in practical applications. For a more detailed comparison, this study has generated histograms depicting the encryption and decryption performance of the four methods, as shown in Fig. 11.Fig. 11a to Fig. 11d present histograms for HE, DP, SMPC, and the research model. The pixel value distributions in HE and DP are relatively concentrated, with noticeable peaks, indicating that significant feature information remains after encryption. In contrast, SMPC displays a more uneven pixel value distribution, with high peaks in certain intervals, suggesting that its encryption effect is not uniform. On the other hand, the research model exhibits a more uniform distribution of pixel values, averaging around 100 feature values, and approaches a smooth distribution. This uniformity indicates that the encrypted image information is more evenly dispersed, effectively concealing feature details and enhancing the effectiveness of privacy protection. However, the method may have the risk of insufficient feature information in extreme cases, which affects the quality of the decrypted image. Future research can explore how to dynamically adjust the encryption strength in different application scenarios to achieve the best balance between privacy protection and image quality. This study evaluates the model using four widely recognized privacy metrics: identity privacy, facial expression privacy, face shape privacy, and posture privacy. Identity privacy is used to evaluate the retention of identity features by the model in protecting facial information. Facial expression privacy is used to evaluate the model's retention of facial expression features. Facial shape privacy is used to evaluate the model's retention of facial features. Posture privacy is used to evaluate the model's retention of attitude features. The test results for these metrics are presented in Tab. 2.

Table 2 Multi-indicator test results for different methods

Data	Model	Retention rate / %				*P*
		Identity privacy	Facial expression privacy	Face shape privacy	Posture privacy	0.003
Figure 1	HE	87.43	88.26	79.64	84.28	0.004
	DP	89.61	86.74	83.19	86.39	0.005
	SMPC	91.27	87.52	85.76	88.72	0.003
	Research model	94.38	89.17	91.28	93.28	0.002
Figure 2	HE	88.68	89.64	87.99	89.37	0.003
	DP	84.51	91.28	83.37	90.26	0.002
	SMPC	90.89	90.36	89.64	91.35	0.003
	Research model	92.54	93.34	91.22	93.57	0.002

In Tab. 2, the privacy retention rates for the two test face images from Fig. 1 and Fig. 2 show that all four methods generally perform well. However, when examining details such as facial expression privacy and posture privacy, the retention rates for the HE, DP, and SMPC methods are relatively moderate. In contrast, the research method demonstrates advantages across all four categories. Quantitative data reveal that the new model achieves the highest retention rates for identity privacy, facial expression privacy, facial shape privacy, and posture privacy, with values of 94.38%, 93.34%, 91.28%, and 93.57%, respectively. These results indicate that the research method is particularly well-suited for current IoT-based personal FIP and has significant practical application value. These results show that the new model has a significant improvement in privacy protection ability and is statistically significant (all P-values < 0.05). This indicates that the model is more effective in protecting personal facial information in the IoT environment and has a strong potential for practical application.

5 Conclusion

This study addressed the challenge of FIP in IoT by proposing a novel model that combines DCGAN and TPM. Compared with the existing methods, the model showed excellent performance in terms of encryption and decryption accuracy, speed, resource efficiency, and privacy protection. The model achieved a balance between privacy preservation and image quality by providing more than 90% privacy protection for key facial attributes. These results showed that the model had great potential for practical applications of FIP in the IoT. The study not only achieved an effective balance between image quality and security but also provided an innovative and practical solution for personal information protection in IoT environments, which is of great practical application value. Nevertheless, the model's scope is constrained by the absence of comprehensive assessments of its performance in extreme environments and in the context of sophisticated cyber-attacks. This limitation may potentially impact the model's reliability in practical applications. Future research should focus on enhancing the robustness of the model to various cyber-attacks and improving its ability to generalize across different datasets.

6 Acknowledgments

This work was supported by Inner Mongolia Autonomous Region Basic Scientific Research Expenses Project for Affiliated Universities and Inner Mongolia Agricultural University Youth Teachers' Scientific Research Ability Improvement Program under grant no.BR230211 and no.BR230210.

7 REFERENCE

[1] Yan, C., Meng, L., Li, L., Zhang, J. H., Wang, Z., Yin, J., Zhang, J. Y., Sun, Y., & Zheng, B.（2022）. Age-invariant face recognition by multi-feature fusion and decomposition with self-attention. ACM Transactions on Multimedia Computing. *Communications, and Applications* (*TOMM*), *18*(1). 1-18. https://doi.org/10.1145/3472810

[2] Saraswat, A. K. & Meel, V. (2022). Protecting data in the 21st century: Challenges, strategies and future prospects. *Information technology in industry*, *10*(2), 26-35.

[3] Hacker, P. (2023). Manipulation by algorithms. Exploring the triangle of unfair commercial practice, data protection, and privacy law. *European Law Journal*, *29*(1-2), 142-175. https://doi.org/10.1111/eulj.12389

[4] Ravi, S., Climent-Pérez, P., & Florez-Revuelta, F. (2024). A review on visual privacy preservation techniques for active and assisted living. *Multimedia Tools and Applications*, *83*(5), 14715-14755. https://doi.org/10.1007/s11042-023-15775-2

[5] Sun, L., Chen, J., & Xu, Y. (2022). Hierarchical amortized GAN for 3D high resolution medical image synthesis. *IEEE journal of biomedical and health informatics*, *26*(8), 3966-3975. https://doi.org/10.1109/JBHI.2022.3172976

[6] Xie, Y., Li, P., Nedjah, N., Gupta, B. B., Taniar, D., & Zhang, J. (2023). Privacy protection framework for face recognition in edge-based Internet of Things. *Cluster Computing*, *26*(5), 3017-3035. https://doi.org/10.1007/s10586-022-03808-8

[7] Sardar, A., Umer, S., Rout, R. K., & Khan, M. K. (2022). A secure and efficient biometric template protection scheme for palmprint recognition system. *IEEE Transactions on Artificial Intelligence*, *4*(5), 1051-1063. https://doi.org/10.1109/TAI.2022.3188596

[8] Dhinakaran, D. & Prathap, P. M. J. (2022). Protection of data privacy from vulnerability using two-fish technique with Apriori algorithm in data mining. The Journal of Supercomputing, *78*(16), 17559-17593. https://doi.org/10.1007/s11227-022-04517-0

[9] Fauzan, F., Lubis, A. F., & Febryani, E. (2024). Analysis of the Influence of Freedom of Expression, Access to Information, and Political Participation on the Protection of Political Rights. *West Science Law and Human Rights*, *2*(01), 62-69. https://doi.org/10.58812/wslhr.v2i01.602

[10] Hasan, M. R., Guest, R., & Deravi, F. (2023). Presentation-level privacy protection techniques for automated face recognition-A survey. *ACM Computing Surveys*, *55*(13), 1-27. https://doi.org/10.1145/3583135

[11] You, A., Kim, J. K., Ryu, I. H., & Yoo, T. K. (2022). Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. Eye and Vision, *9*(1), 6-7. https://doi.org/10.1186/s40662-022-00277-3

[12] Wang, Y., Yan. X., Guan, D., Wei, M., Chen, Y., Zhang, X. P., & Li, J. (2022). Cycle-snspgan: Towards real-world image dehazing via cycle spectral normalized soft likelihood estimation patch gan. *IEEE Transactions on Intelligent Transportation Systems*, *23*(11), 20368-20382. https://doi.org/10.1109/TITS.2022.3170328

[13] Mahmoudinejad., Dezfuli, S. R., & Kheyrandish, M. (2022). An image encryption method based on chaotic system exploiting fuzzy system and arithmetic coding. *Multimedia Tools and Applications*, *81*(30), 44263-44289. http://dx.doi.org/10.1007/s11042-022-13250-y

[14] Chen, B., Tan, W., Wang, Y., & Zhao, G. (2022). Distinguishing between natural and GAN-generated face

images by combining global and local features. Chinese *Journal of Electronics*, *31*(1), 59-67.

[15] Raposo, V. L. (2023). The use of facial recognition technology by law enforcement in Europe: a non-orwellian draft proposal. *European Journal on Criminal Policy and Research*, *29*(4), 515-533. https://doi.org/10.1007/s10610-022-09512-y

[16] Akraam, M., Rashid, T., & Zafar, S. (2023). An image encryption scheme proposed by modifying chaotic tent map using fuzzy numbers. *Multimedia Tools and Applications*, *82*(11), 16861-16879. https://doi.org/10.1007/s11042-022-13941-6

[17] Peng, F., Yin, L., & Long, M. (2022). BDC-GAN: Bidirectional conversion between computer-generated and natural facial images for anti-forensics. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(10), 6657-6670. https://doi.org/10.1109/TCSVT.2022.3177238

[18] Almomani, I., El-Shafai, W., & Al Khayer A. (2023). Proposed biometric security system based on deep learning and chaos algorithms. Comput. Mater. *Contin*, *74*(2), 3515-3537. https://doi.org/10.32604/cmc.2023.033765

[19] Medin, S. C., Egger, B., Cherian, A., Wang, Y, & Marks, T. K. (2022). MOST-GAN: 3D morphable Style GAN for disentangled face image manipulation. *Proceedings of the AAAI conference on artificial intelligence*, *36*(2), 1962-1971. https://doi.org/10.1609/aaai.v36i2.20091

[20] Huang, G. & Jafari, A. H. (2023). Enhanced balancing GAN: Minority-class image generation. *Neural computing and applications*, *35*(7), 5145-5154. https://doi.org/10.1007/s00521-021-06163-8

[21] Wen, H., Xie, Z., Wu, Z., Lin, Y., & Feng W. (2024). Exploring the future application of UAVs: Face image privacy protection scheme based on chaos and DNA cryptography. *Journal of King Saud University-Computer and Information Sciences*, *36*(1), 101871. https://doi.org/10.1016/j.jksuci.2023.101871

[22] Xu, S., Chang, C. C., & Nguyen, H. H. (2024). Reversible anonymization for privacy of facial biometrics via cyclic learning. *EURASIP Journal on Information Security*, (1), 24-27. https://doi.org/10.1186/s13635-024-00174-3

[23] Kováč, P., Jackuliak, P., & Bražinová, A. (2024). Artificial Intelligence-Driven Facial Image Analysis for the Early Detection of Rare Diseases: Legal, Ethical, Forensic, and Cybersecurity Considerations. *AI*, *5*(3), 990-1010. https://doi.org/10.3390/ai5030049

[24] Atanasov, I. & Pilev, D. (2024). Cyber-Physical Security Through Facial Recognition And Sensor Data Analysis. *Journal of Chemical Technology and Metallurgy*, *59*(2), 465-472. https://doi.org/10.59957/jctm.v59.i2.2024.27

[25] Aloisi, A. (2024). Regulating algorithmic management at work in the European Union: Data protection, non-discrimination and collective rights. *International Journal of Comparative Labour Law and Industrial Relations*, *40*(1), 37-70. httpsl://doi.org/10.54648/ijcl2024001

[26] Cai, Z., Gao, Z., & Planche, B. (2024). Disguise without disruption: Utility-preserving face de-identification. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(2), 918-926. https://doi.org/10.1609/aaai.v38i2.27851

Contact information:

Ye QIN
Vocational and Technical College,
Inner Mongolia Agricultural University,
010018, China
E-mail: nmgndqinye@126.com

Yaowu KANG
(Corresponding author)
Vocational and Technical College,
Inner Mongolia Agricultural University,
010018, China
E-mail: kyw_94@126.com