

Research Article

A Privacy-Preserving Biometric Recognition System with Visual Cryptography

Lijing Ren ¹ and Denghui Zhang ²

¹School of Traffic and Transportation, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China

²Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China

Correspondence should be addressed to Denghui Zhang; zhang.denghui@foxmail.com

Received 29 December 2021; Revised 8 February 2022; Accepted 18 February 2022; Published 22 March 2022

Academic Editor: Mohammad R. Khosravi

Copyright © 2022 Lijing Ren and Denghui Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The popularity of more powerful and smarter digital devices has improved the quality of life and poses new challenges to the privacy protection of personal information. In this paper, we propose a biometric recognition system with visual cryptography, which preserves the privacy of biometric features by storing biometric features in separate databases. Visual cryptography combines perfect ciphers and secret sharing in cryptography with images, thus eliminating the complex operations in existing privacy-preserving schemes based on cryptography or watermarking. Since shares do not reveal any feature about biometric information, we can efficiently transmit sensitive information among sensors and smart devices in plain. To abate the influence of noise in visual cryptography, we leverage the generalization ability of transfer learning to train a visual cryptography-based recognition network. Experimental results show that our proposed method keeps the high accuracy of the feature recognition system when providing security.

1. Introduction

With the help of smartphones equipped with rich sensors and high-bandwidth 5G networks and beyond [1], now we can share information at any time and any place [2]. Face images have been the most widely used perceptual information transmitted on the Internet [3, 4]. With people paying more attention to the privacy of personal information, it has been a major concern for ensuring privacy in these crucial services provided on public networks [5]. Traditional cryptography methods based on a password or ID card have shortcomings including easy to forge, easy to forget, and computation complexity, which prevents them from widespread applications [6].

Visual cryptography (VC) [7, 8] (also known as k -out-of- n VC) is a secret-sharing method aiming at images. It splits a secret image into n shares. The threshold characteristic makes it impossible to restore the secret image unless stacking k , ($k < n$) or more shares together. The human visual system (HVS) can recover secret information by simply printing shares on transparencies and stacking them,

without any digital device. VC provides a simple and effective method for the distributed storage of feature data, where there is no need to maintain keys in encryption. These features of VC are particularly suitable for scenarios in a limited-computing and untrusted networking environment. VC eliminates the complex computation required by traditional watermarking or cryptography. However, recovery images from a VC scheme (VCS) have poor quality and expanding size [9].

In this paper, we propose a novel VC-based privacy-preserving method during biometric recognition [10]. This paper first constructs a VCS to avoid noise-like and expanding shares. Then, we distribute and store images with VCS. Last, we use the transfer learning method [11] to learn features from separated databases. Our contributions of this paper are as follows:

- (1) We construct a meaning VCS, where HVS can print shares on transparencies and recover secret images
- (2) We propose a secure storage method for biometric features by the secret sharing of images

- (3) We keep the recognition performance of images recovered from VC with the strong generalization of transfer learning

The structure of the rest of the paper is as follows. Section 2 provides helpful background on feature recognition and improved VCSs. Section 3 describes our privacy-preserving recognition model. Section 4 evaluates the performance of our approach on two classic datasets. Finally, we summarize our contributions in Section 5.

2. Related Work

2.1. Feature Recognition. The past decades have witnessed the explosion of advances in feature recognition [12]. By collecting massive samples with big data technologies, deep neural networks (DNN) [13, 14] can use an artificial neural network to directly give inference results.

Metric learning [15] is an emerging feature recognition method that can verify whether two embedding samples belong to the same identity or not. Authors [16] use a DNN to transform a face image into a vector and then calculate and compare the Euclidean distance between two vectors. Directly using Euclidean distance is equivalent to considering the only intraclass distance. However, sometimes the intraclass distance may be larger than the interclass distance. To address the shortcomings of Euclidean distance, FaceNet [17] used a triplet-based loss function to embed face images. Minimizing the triplet loss function is both minimizing the distance of similar samples and maximizing the distance of nonsimilar samples at the same time.

To train a model enhancing the discrimination of learned features, Wen et al. [18] join supervision of center loss and softmax cross-entropy loss as the loss function of neural networks, thus significantly accelerating the convergence of training models.

2.2. Visual Cryptography. VC is a (k, n) threshold scheme for images. The simple decryption makes VC surpass other cryptographic schemes based on cryptography or watermark [19]. VC expands the application scenarios of secret sharing. It has been an emerging research field in the field of image cryptography since born in 1995 [9, 20, 21]. The conditional VCS encrypts a secret image into shares pixel-by-pixel, hence resulting in pixel expansion. It can eliminate pixel expansion by mapping a block in the secret image into an equal-sized block at the corresponding position of shares. Through vertical arrangements and elaborately processing the correspondence between the secret block and share blocks, we can keep the size and obtain a higher contrast than previous schemes. Another scheme keeping the size of recovery images is the probabilistic VCS which firstly randomly selects a column from a basis matrix and then distributes each pixel in the column to the corresponding position of shares.

The EVCS is also known as friendly VCS[22], which issues the management difficulty caused by noise-like shares in VCS. Ateni [23] proposed a general technique to implement EVCS for any access structure by hypergraph coloring. Due to the pixel-by-pixel encryption mechanism,

restored images remain the problem of pixel expansion. Lee [24] proposed a novel algorithm of general access structures to cope with the pixel expansion problem. Instead of the traditional VC-based approach, this scheme first constructed meaningless shares using an optimization technique and then solved it by a simulated-annealing algorithm. Then, the method generates meaningful shares by adding cover images into shares with a stamping algorithm. No computational devices are needed for decryption in this scheme. However, it is only applicable to black-and-white images and needs a lot of time to generate shares.

To address the risk of leakage in biometric features, Jinu [25] proposed a multifactor authentication scheme based on VC and Siamese networks. However, the VCS adopted needs a key to encrypt, which destroys the printable characteristic of VC. Ross [26] preserved the privacy of face template data through a trusted third party and EVCS. In this scheme, shares for decoupling secret face data come from a group of general face images. To improve the quality of restored images, it may use up to 100 shares to encrypt a face image. How to transmit these shares will be a large challenge. Requiring huge storage blocks its application in many scenarios. The accuracy of face recognition decreases because of the interference of VC or EVC in these schemes.

3. A Privacy-Preserving Biometric Recognition System with Visual Cryptography

Combining feature identification and cryptography can effectively build a secure feature recognition system. In this section, we first present our novel EVCS to address the pixel expansion and vulnerability of noise-like shares, and then we use the proposed scheme to securely distribute face images in separate databases. In the end, to keep the accuracy of face recognition, we leverage the transfer learning method to mitigate the quality degradation of recovery images.

3.1. Expansion-Free EVCS. In the traditional pixel-by-pixel encryption of VC, there is a corresponding matrix collection C_0 and C_1 , for a white (w) or black (b) pixel in a secret image. The matrix in C_0 or C_1 consists of $n \times m$ Boolean values. When encrypting, we randomly select a matrix from C_0 or C_1 and assign n rows of pixels of the matrix to the corresponding n shares. Each row contains m subpixels which are reinterpreted as recovery w or b pixel.

When superimposing enough shares, the secret pixel becomes visible. The gray-level of a subpixel is proportional to the Hamming weight ($H(V)$), which denotes the number of black pixels in a combined vector V from basis matrices S^0 and S^1 . The HVS interprets the recovery pixels as black if $H(V) \geq d$ and as white if $H(V) < d - am$, where a (contrast) is the difference of $H(V)$ between the white pixel and black pixel in the recovery image and m is the pixel expansion. The a and m are two important parameters that govern the quality of reconstructed images. For a VCS, we would like a to be as large as possible, and m as small as possible (close to 1). A VCS has to satisfy the following three conditions to keep security and validity:

- (i) For any matrix S^0 in C_0 , $H(V) < d - am$ is satisfied when overlaying any k of n row vectors in S^0
- (ii) For any matrix S^1 in C_1 , $H(V) \geq d$ is satisfied when overlaying any k of n row vectors in S^1
- (iii) For any subset with q , $q < k$ rows, rows overlaying of $q \times m$ matrices are indistinguishable, which implies the black-and-white pixels in combined shares have a uniform probability distribution

Conditions (i) and (ii) make images visible when overlaying. The condition (iii) implies that fewer than k shares cannot gain any information about the secret pixel.

Taking the (2, 2)-VCS with two participants for example, if the pixel arrangement held by the other participant is complementary to itself, all black pixels will appear, while if pixels held by the other are the same as itself, it will generate half black and half white gray pixels after superimposing. The proportion of black-and-white pixels in each share is fixed, so it will not reveal any information about a secret image.

The disadvantage of VC is that generated shares are meaningless noisy-like images. Although such images will not reveal any information about secret images, they will increase the burden of management and pique the interest of attackers. If these images are tempered by malicious attackers, it is difficult to detect.

The difference between VCS and EVCS is that EVCS has to take into account the color of shares besides the secret image we want to get [27], hence we can see $n + 1$ meaningful images in an EVCS. Table 1 shows collections for traditional (2,2)-EVCS and recovered subpixels, where the gray-level 1(0) denotes a white (black) pixel (w, b). Each secret pixel is expanded to 4 subpixels in the EVCS ($m = 4$). Only when enough participants show held shares can the secret images be restored. For example, if the pixel is black in the secret image, black in *Share*₁, and white in *Share*₂, the share subpixels will be selected from the third row and second column of Table 1, that is, the [0 1 0 0] and [1 0 1 0] subpixel blocks. After superimposing the two rows, we get a subpixel with four black pixels. The subpixel blocks [0 1 0 0] and [1 0 1 0] may also appear in other basis matrices (e.g., S_b^{bb} and S_w^{ww}), hence the secret information is not exposed in the cover images. When encrypting a black secret pixel, regardless of which basis matrix is selected to generate the corresponding subpixels, we can obtain a whole black subpixel block ([0 0 0 0]), while for white secret pixel, we can only obtain a subpixel block with three black pixels and one white pixel ([1 0 0 0]), which is interpreted as a white pixel in the recovery image, hence the recovery pixel shows a contrast of 1/4 (three white pixels are lost). It can also be verified that for the (2,2)-EVCS, the contrast values of the covers image also are $a = 1/4$. No matter which basic matrix we choose, the white pixel in *Share*₁ or *Share*₂ is represented by a subpixel with two black pixels and two white pixels, while the black pixel is represented by a subpixel with three black white pixels and one black pixel, so the contrast in cover images is $2/4 - 1/4 = 1/4$.

To eliminate the pixel expansion problem in the traditional EVCS model, we can use the block-wise operation instead of pixel-by-pixel encryption [28]. We first divide a

gray-level image into n nonoverlapping black-and-white pixel blocks B_i , $B_i \cap B_j = \emptyset$, for $1 \leq i \neq j \leq n$. B_i and half-toned blocks B_n are the same in size. The number of black pixels in B_i and B_n must satisfy the following: $b_{Bi} = s_i / (s_b + 1) \times (s_1 - \sum B_i / 255)$, where s_i denotes the size of candidate black levels, b_{Bi} denotes the gray-level of cover blocks, and s_b denotes the block size. To meet the security requirement of VC, the contrast of black-and-white blocks cannot be any, which results in contrast loss after halftoning.

We reinterpret the image blocks as white and black pixels based on the original EVCS. For example, for the (2,2)-EVCS, we will dither a secret gray-level block into the closest equal-sized pixel block with 3 or 4 black pixels and dither a cover block into the closest equal-sized pixel block with 2 or 3 black pixels. If the encryption block has three black pixels (e.g., [1 0 0 0]) and cover blocks have two black (e.g., [0 1 1 0]) and three black pixels, e.g., [0 1 0 0], respectively, then we will set c_s to w and c_{h1}, c_{h2} to w, b . Therefore, we will select S_w^{wb} as the basis matrix and randomly permuting all columns of S_w^{wb} to obtain the collections C . At last, we select a matrix C_p from C and dispatch each row of C_p to the corresponding blocks in share, $i \in [i \in 1, \dots, n]$.

After processing the secret image by the limited halftoning, we can reuse the underlying EVCS directly to encrypt the secret image into meaningful share images without pixel expansion. Ateni [23] et al. describe a general construction method for EVCS in detail.

3.2. The Recognition of Face Images Recovered from EVC.

In this section, we will propose a new embedding method for VC-recovered images to maintain high accuracy in feature identification. Although the proposed EVCS solves the problems of size expansion and noise-like shares, it still cannot achieve perfect image restoration. For existing feature identification methods, the use of VC can alleviate the leakage of centralized storage of template data, but it will also mix noise into sample images, which will reduce the performance of feature recognition.

Witnessing the recent success of DNN, we hope to use these methods to solve the problem of noise interference caused by the introduction of VC in feature identification [6]. Figure 1 shows the flowchart of the proposed approach for distributing and matching face images. In a high-precision neural network model for face images, the face data is converted into the corresponding weights in the network. Using transfer learning [29], we can extract and transfer these weights to other neural networks (e.g., reduced images mixed with noisy signals). Transfer learning allows sharing the learned model parameters and structures to a new model in a specific way, thus speeding up and optimizing the learning efficiency of the model and avoiding learning from scratch. We first train a softmax classifier on the training data using a pretrained neural network model and then fine-tune the weights of the last layer or layers using a dataset recovered from EVCS. To avoid affecting learned weighted when training, we freeze the pretrained model and add multiple fully connected layers at the end of the network.

TABLE 1: The collections for (2,2)-EVCS example and the recovered subpixels.

	Secret pixels		Recovery subpixels	
	White pixel (1)	Black pixel (0)	White pixel	Black pixel
Basis matrix	$S_w^{ww} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$	$S_b^{ww} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
	$S_w^{wb} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	$S_b^{wb} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$		
	$S_w^{bw} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$	$S_b^{bw} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$		
	$S_w^{bb} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	$S_b^{bb} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$		

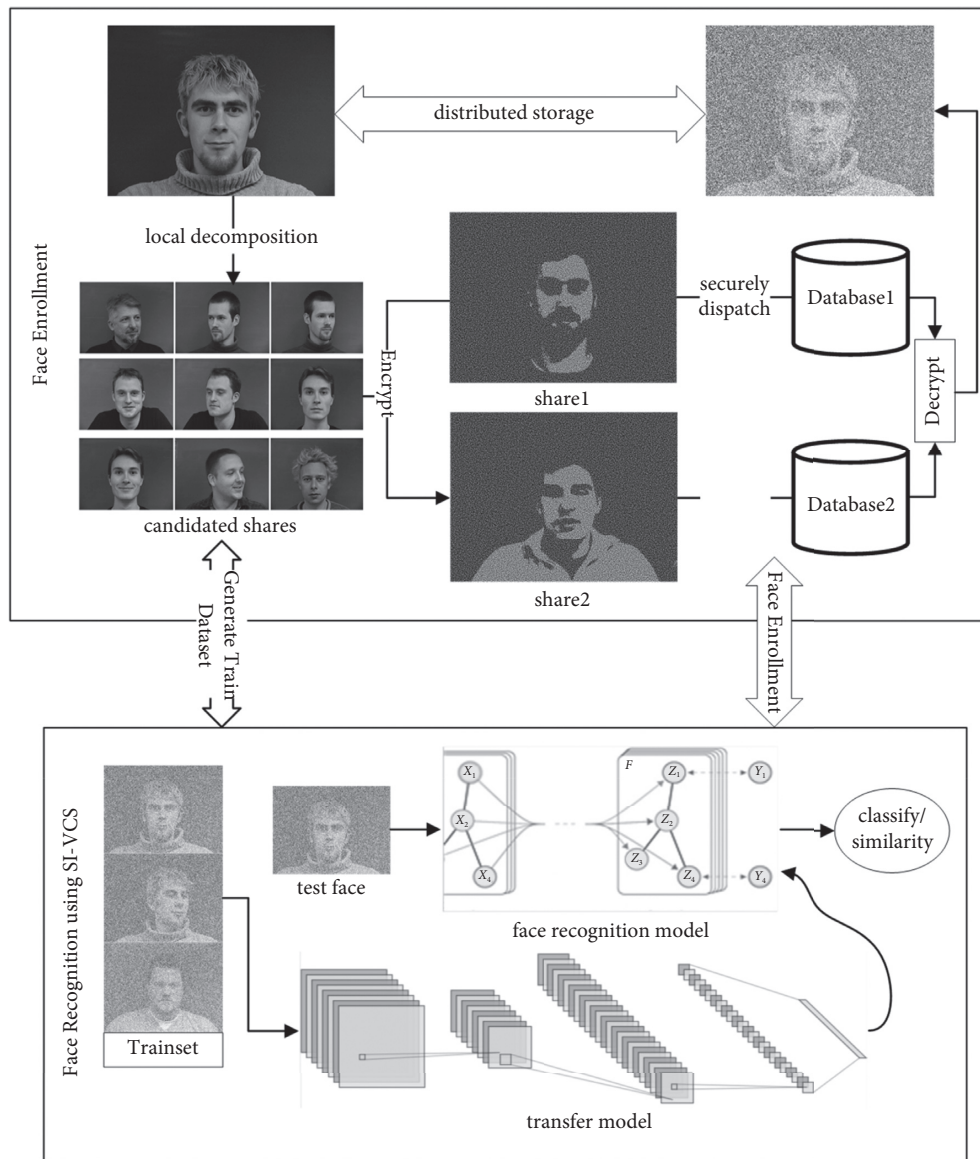


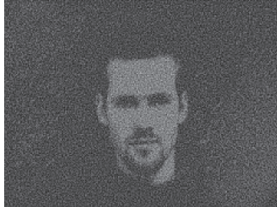



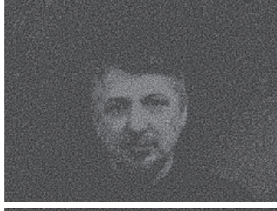


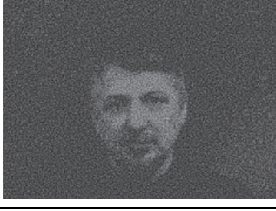




FIGURE 1: The flowchart of the proposed approach for distributing and matching face images.

After embedding reduced images, we select a loss function to efficiently obtain a high-precision model. When used as a classification task, the model selects the most similar image class ID in the training set to the test image as the output.

In deep learning, many methods use pairs of samples to compute gradient loss and update model parameters [15]. We use the triplet loss [17] gradient optimization in the face recognition system. Although center loss [18] may have a

TABLE 2: The decomposed shares and decrypted images for one embedding.

Type	Secret image	Share1	Share2	Decrypted image
Anchor				
Positive				
Negative				

faster convergence rate, triplet loss is more robust for image quality degradation. The basic idea of triplet loss is to make the distance between negative sample pairs larger than that between positive sample pairs. In the training process, positive sample pairs and negative sample pairs are selected at the same time, and the anchor of the positive and negative sample pairs is the same. Choosing which triplets to train is important for achieving good performance. When the distance between the negative sample and the positive sample pairs is greater than a threshold, the model sets the loss to 0 and ignores the sample pairs. Because the distance between positive and negative samples and anchor points is considered at the same time, the performance using triplet loss is often better than that of contrastive loss [30].

To protect the privacy of sensitive biometrics, we decompose the original image into two or more meaningful images. The similarity calculation of samples is realized as

$$L = \max(d(a, p) - d(a, n) + \text{margin}, 0). \quad (1)$$

The triplet loss minimizes L , then $d(a, p) \rightarrow 0, d(a, n) \rightarrow \text{margin}$. The goal is to shorten the distance between a and p and lengthen the distance between a and n . For the case $L = 0$, that is, $d(a, p) + \text{margin} < d(a, n)$, this is an easily discernible triple and does not need to optimize, because a and p and a and n are far apart; for the case $d(a, p) < d(a, n) < d(a, p) + \text{margin}$, the triples are in the fuzzy region. This case is the focus of our training.

The same deep neural network is used to extract the features of these three images to obtain three embedding vectors, which are input parameters of triple loss. The model is updated by using a backpropagation algorithm according to loss and iterated until a stable neural network model is obtained. When inferring samples, we first generate triplet

image pairs and then output the distances between anchor, positive, and negative embedding [31]. After transforming the distances into probability models, we can get desired results. During the registration process, we decompose the private face data into two or more meaningful images. After dispatching shares, the system discards the original data. The encrypted face data are stored in two or more database servers. Unless these servers collude, the private data will not be revealed to any server. In the authentication stage, the feature recognition system sends requests to database servers to transmit corresponding shares to it. Once finishing the classification recognition or similarity matching task, our system will discard the reconstructed secret image. In the whole registration and recognition process, the secret image will be only recovered in use. Because private biometrics cannot be extracted from any single database or server, the whole authentication does not reveal any feature information.

4. Experiments and Results

In this section, we first evaluate the effectiveness of the proposed method. We select the TensorFlow [32] framework to develop the Siamese network [31] model and the test application using Python rapidly.

Table 2 shows a triple example in the IMM dataset [33]. Every embedding has three input images, which are anchor, positive, and negative images, respectively. The proposed EVCS keeps the image size before and after encryption. All images are 640×480 in size.

Figure 2 shows the encryption time comparison between our and Lee's [24] schemes for three images from the IMM dataset. When encrypting one image, we selected two other images separately as cover images. It can be seen that Lee's method took a lot of time in the encryption process, which

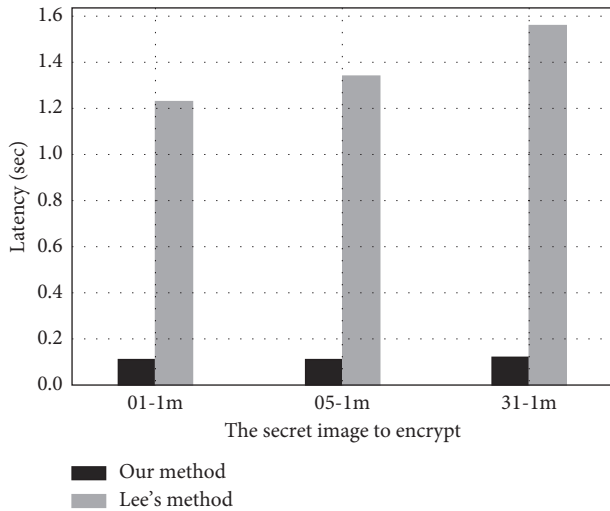


FIGURE 2: The encryption time comparison between our and Lee’s [24] schemes.

TABLE 3: The performance of our networks on IMM and LFW datasets for normal and recovery images, respectively.

Accuracy	IMM	Normal	0.97
		Our	0.95
	LFW	Normal	0.95
		Our	0.93
Precision	IMM	Normal	0.97
		Our	0.96
	LFW	Normal	0.97
		Our	0.93
Recall	IMM	Normal	0.98
		Our	0.92
	LFW	Normal	0.92
		Our	0.90
F1-score	IMM	Normal	0.98
		Our	0.94
	LFW	Normal	0.93
		Our	0.92

was 10X that of our method. It is because Lee’s method developed a complex simulated-annealing-based method to generate an optimized arrangement when encrypting the cover images, while our algorithm can directly obtain the encrypted pixels for each share from the base matrix. So, the encryption performance of our algorithm is better.

We further test the recognition system on the IMM and LFW datasets. We choose ResNet18 [14] as the backbone network and use pretrained weights from the ImageNet [34] dataset. The IMM Face Database comprises 240 still images of 40 different human faces. LFW (Labeled Faces in the Wild) [35] is the de-facto benchmark for face verification, also known as pair matching. The dataset contains more than 13,000 labeled images of faces collected from the web. 1680 of the people pictured have two or more distinct photos in the dataset.

Table 3 shows the performance including accuracy, precision, recall, and f1-score of our networks on IMM and LFW datasets for normal and images recovered from the

proposed EVCS, respectively. Because the normal data is unencrypted, the recognition performance is higher. For the dataset recovered from the proposed EVCS, the recognition performance is lower than that of normal data because the image is lossy. However, this experiment shows the trained model can archive results of more than 0.92 on all metrics, which implies we can still keep the inferred performance of the trained model when securing face images. The reason the metrics on the LFW dataset are lower than IMM is that the face consists of more complex patterns. The proposal processes every pixel in the face image one by one while preserving the gray-level density. This transformation corresponds to adding a global white noise to the image, which does not destroy the features and patterns in face images, so it does not affect the final prediction accuracy.

5. Conclusions

With the development of next-generation networks and smart devices with rich sensors, it is convenient to sense and transmit images anytime and anywhere. There is an urgent need to protect the security and privacy of easily collected information and efficiently transfer them in untrusted networks. In this paper, we propose a novel feature recognition method taking advantage of VC and use transfer learning to improve the recognition system. This proposed method eliminates the complex computation in traditional cryptography. To solve the problems of pixel expansion and noise-like shares in VC, we propose a novel EVCS to keep the size of recovered images by block encryption before and after encryption. We further utilize the strong generalization ability of transfer learning to eliminate the interference of noise for images recovered from EVCS. The experimental results show our method can keep the high accuracy of feature recognition when preserving privacy.

In future work, we will combine VC with other methods like QR codes to provide richer forms for the transmission of private images. We will also test the performance of the proposed method in other biometrics such as fingerprints and voiceprints to verify its popularity.

Data Availability

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

Conflicts of Interest

All authors have no conflicts of interest to declare.

Acknowledgments

This work is supported in part by the National Key Research and Development Program of China (2019YFB1706003), and the Natural Science Foundation of Guangdong Province, China (2214050004382).

References

- [1] Z. Tian, S. Su, W. Shi, X. Du, M. Guizani, and X. Yu, “A data-driven method for future Internet route decision modeling,”

- Future Generation Computer Systems*, vol. 95, pp. 212–220, 2019.
- [2] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, “A survey on internet of things: architecture, enabling technologies, security and privacy, and applications,” *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, 2017.
 - [3] X. Jin, Y. Li, S. Ge, C. Song, L. Wu, and X. Zhou, “Secure face retrieval for group mobile users,” *Soft Computing*, vol. 23, no. 23, Article ID 12820, 2019.
 - [4] Z. Gu, L. Wang, X. Chen et al., “Epidemic risk assessment by A novel communication station based method,” *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 1, pp. 332–344, 2022.
 - [5] Z. Gu, H. Li, S. Khan et al., “IEPSBP: a cost-efficient image encryption algorithm based on parallel chaotic system for green IoT,” *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 89–106, 2022.
 - [6] N. Akhtar and A. Mian, “Threat of adversarial attacks on deep learning in computer vision: a survey,” *IEEE Access*, vol. 6, Article ID 14430, 2018.
 - [7] M. Naor and A. Shamir, “Visual cryptography,” in *Proceedings of the Advances in Cryptology - EUROCRYPT’94*, pp. 1–12, Springer, Perugia, Italy, May 1994.
 - [8] D. Zhang, H. Zhu, S. Liu, and X. Wei, “HP-VCS: a high-quality and printer-friendly visual cryptography scheme,” *Journal of Visual Communication and Image Representation*, vol. 78, Article ID 103186, 2021.
 - [9] J. Weir and W. Yan, “A comprehensive study of visual cryptography,” in *Transactions on Data Hiding and Multimedia Security V*, pp. 70–105, Springer, NY, USA, 2010.
 - [10] S. Kumar, S. K. Singh, A. K. Singh, S. Tiwari, and R. S. Singh, “Privacy preserving security using biometrics in cloud computing,” *Multimedia Tools and Applications*, vol. 77, no. 9, Article ID 11039, 2018.
 - [11] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, “Momentum contrast for unsupervised visual representation learning,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9729–9738, Seattle, WA, USA, 13–19 June 2020.
 - [12] K. Sundararajan and D. L. Woodard, “Deep learning for biometrics,” *ACM Computing Surveys*, vol. 51, no. 3, pp. 1–34, 2019.
 - [13] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, “DeepFace: closing the gap to human-level performance in face verification,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1701–1708, IEEE, Columbus, OH, USA, June 2014.
 - [14] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
 - [15] M. Norouzi, D. J. Fleet, and R. R. Salakhutdinov, “Hamming distance metric learning,” in *Advances in Neural Information Processing Systems*, pp. 1061–1069, Cambridge, MA, United States, 2012.
 - [16] C. K. Hsieh, L. Yang, Y. Cui, T. Y. Lin, S. Belongie, and D. Estrin, “Collaborative metric learning,” in *Proceedings of the 26th International Conference on World Wide Web*, pp. 193–201, Perth Australia, April 2017.
 - [17] F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: a unified embedding for face recognition and clustering,” in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, Boston, MA, USA, June 2015.
 - [18] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, “A discriminative feature learning approach for deep face recognition,” in *Proceedings of the Computer Vision - ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., pp. 499–515, Cham: Springer International Publishing, Amsterdam, The Netherlands, October 2016.
 - [19] T. M. Thanh and K. Tanaka, “An image zero-watermarking algorithm based on the encryption of visual map feature with watermark information,” *Multimedia Tools and Applications*, vol. 76, no. 11, Article ID 13471, 2017.
 - [20] M. Naor, A. Shamir, and A. Shamir, “Visual Cryptography II: Improving the Contrast via the Cover Base,” in *Proceedings of the Security Protocols*, Paris, France, April 1997.
 - [21] Y. Cheng, Z. Fu, and B. Yu, “Improved visual secret sharing scheme for QR code applications,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2393–2403, 2018.
 - [22] I. InKoo Kang, G. R. Arce, and H.-K. Heung-Kyu Lee, “Color extended visual cryptography using error diffusion,” *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 132–145, 2011.
 - [23] G. Ateniese, C. Blundo, A. De Santis, and D. R. Stinson, “Extended capabilities for visual cryptography,” *Theoretical Computer Science*, vol. 250, no. 1–2, pp. 143–161, 2001.
 - [24] K.-H. Lee and P.-L. Chiu, “An extended visual cryptography algorithm for general access structures,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 219–229, 2012.
 - [25] J. Mohan and R. Rajesh, “ENHANCING home security through visual CRYPTOGRAPHY,” *Microprocessors and Microsystems*, vol. 80, Article ID 103355, 2021.
 - [26] A. Ross and A. Othman, “Visual cryptography for biometric privacy,” *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 70–81, 2010.
 - [27] M. Nakajima and Y. Yamaguchi, “Extended visual cryptography for natural images,” *Journal of WSCG*, vol. 10, no. 2, pp. 303–310, 2002.
 - [28] Y.-C. Hou, Z.-Y. Quan, C.-F. Tsai, and A.-Y. Tseng, “Block-based progressive visual secret sharing,” *Information Sciences*, vol. 233, pp. 290–304, 2013.
 - [29] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
 - [30] R. Hadsell, S. Chopra, and Y. LeCun, “Dimensionality reduction by learning an invariant mapping,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 1735–1742, 2006.
 - [31] H. Pang, Q. Xuan, M. Xie, C. Liu, and Z. Li, “Research on Target Tracking Algorithm Based on Siamese Neural Network,” *Mobile Information Systems*, vol. 2021, Article ID e6645629, 11 pages, 2021.
 - [32] M. Abadi, P. Barham, J. Chen, Z. Chen, and X. Zhang, *TensorFlow: A System for Large-Scale Machine Learning*, USENIX Association, Berkeley, California, United States, 2016.
 - [33] M. M. Nordström, M. Larsen, J. Sierakowski, and M. B. Stegmann, “The IMM Face Database - an Annotated Dataset of 240 Face Images,” Technical University of Denmark, Lyngby, Denmark, Report, 2004.
 - [34] O. Russakovsky, J. Deng, H. Su et al., “ImageNet large scale visual recognition challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
 - [35] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, “Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments,” 2008.