Deep Hashing Based Cancelable Multi-Biometric Template Protection

Guichuan Zhao, Qi Jiang, Ding Wang, Xindi Ma, and Xinghua Li, *Member, IEEE*

***Abstract*—The increasing use of multi-biometric authentication has raised concerns about the security of biometric templates. Many template protection methods based on convolutional neural network have been presented, but most involve a trade-off between authentication accuracy and template security. In this paper, we present a cancelable multi-biometric template protection scheme that combines deep hashing with cancelable distance-preserving encryption (CDPE), which provides high template security without degrading the authentication performance. Specifically, a deep hashing based architecture that minimizes the quantization loss is designed to map face and iris traits to binary codes. Next, CDPE is proposed to generate a protected template given the face binary code and a user-specific key obtained from the iris binary code, which preserves the distance between original templates in the protected domain to ensure authentication performance equivalent to unprotected systems. Digital lockers instead of the key are stored to further enhance the security, which can be unlocked with genuine biometric traits to get the correct key during authentication. Theoretical and experimental results on real face and iris datasets show that our scheme can achieve equal error rate of 0.23% and genuine accept rate of 97.54%, while guaranteeing irreversibility, revocability and unlinkability of protected templates.**

***Index Terms*—Cancelable biometric, deep hashing, multi-biometric, cancelable distance-preserving encryption, reusable fuzzy extractor, template security.**

Manuscript received 19 September 2022; revised 4 August 2023; accepted 15 November 2023. Date of publication 28 November 2023; date of current version 11 July 2024. This work was supported in part by the Major Research plan of the National Natural Science Foundation of China under Grant 92167203, in part by the National Natural Science Foundation of China under Grants 62072352, 62125205, 62372350, 62232013, and 62072359, in part by the Foundation for Innovative Research Groups of the National Natural Science Foundation of China under Grant 62121001, in part by Scientific Research Program Funded by Shaanxi Provincial Education Department under Grant 20JY016, in part by Fundamental Research Funds for the Central Universities, Innovation Fund of Xidian University, Key Industrial Chain Projects in Shaanxi Province under Grant 2020ZDLGY09-06. *(Corresponding author: Qi Jiang.)*

Guichuan Zhao and Qi Jiang are with the School of Cyber Engineering, Xidian University, Xi'an 710071, China (e-mail: 1078161458@qq.com; jiangqixdu@gmail.com).

Xindi Ma is with the School of Cyber Engineering, Xidian University, Xi'an 710071, China, and also with the Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541000, China (e-mail: xdma1989@gmail.com).

Ding Wang is with the College of Cyber Science, Tianjin Key Laboratory of Network and Data Security Technology, Nankai University, Tianjin 300350, China, and also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China (e-mail: wangding@nankai.edu.cn).

Xinghua Li is with the State Key Laboratory of Integrated Services Networks, and the School of Cyber Engineering, Engineering Research Center of Big data Security, Ministry of Education, Xidian University, Xi'an 710071, China (e-mail: xhli1@mail.xidian.edu.cn).

Digital Object Identifier 10.1109/TDSC.2023.3335961

I. Introduction

MULTI-BIOMETRIC authentication have played an increasingly important role in recent years due to its advantages such as lower error rate, higher accuracy, and larger population coverage. Generally, modern biometric authentication mainly use deep convolutional neural networks (CNNs) to extract discriminative features (commonly referred to as a template), from biometric traits, such as facial images. However, recent researches [1], [2], [3], [4] demonstrate that original biometric traits can be reconstructed from unprotected templates, resulting in the leakage of privacy and the permanent loss of identity. In addition, these reconstructed biometric data may be used by attackers to get unauthorized access to the biometric authentication systems. Therefore, privacy and security issues are of great importance in biometric templates, especially in multi-biometric systems where multiple traits of each user are stored. The European Union General Data Protection Regulation has recognized biometric templates as sensitive data that must be protected [5].

The purpose of biometric template protection (BTP) is to achieve irreversibility, renewability and unlinkability as mentioned in ISO/IEC IS24745:2011 [6]. In terms of irreversibility, it must be computationally difficult to reconstruct the raw biometric trait from the corresponding protected template. For renewability, the generation of a new protected biometric template based on different security parameters should be easy. As for unlinkability, it should be indistinguishable between different protected templates generated by the same user biometric traits. In addition, the fundamental requirement of BTP methods is the low intra-user variability and high inter-user variability of the protected templates [7]. As the low intra-user variability can improve the matching performance (low false rejection rate), while the high inter-user variability can reduce false acceptance rate.

To overcome the above challenge, many CNN-based biometric template protection methods have been proposed, which are commonly divided into two categories: *Non-NN* and *NN-learned* [8]. In the *Non-NN* method [9], [10], [11], a neural network (NN) is used as a feature extractor to extract a biometric vector, which is then submitted to the BTP algorithm to generate a protected template (as shown in Fig. 1(a)). The *NN-learned* method [12], [13], [14] applies a specific NN to learn protected templates directly from biometric traits (as shown in Fig. 1(b)).

Compared with *NN-learned* method, *Non-NN* method is more flexible and can be applied to protect feature vectors generated

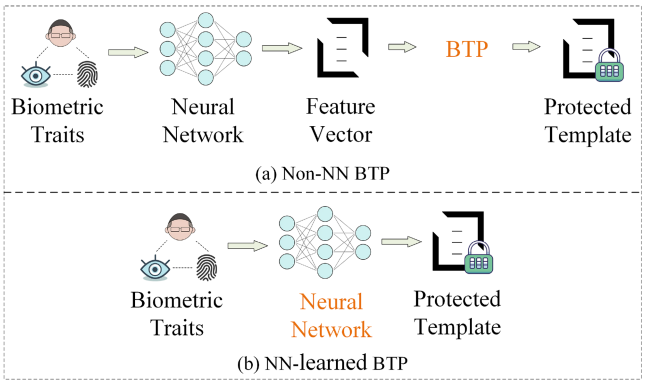


Fig. 1. Non-NN and NN-learned BTP methods.

by other NN models. Besides, a more deterministic definition of the template transformation algorithm in the *Non-NN* method allows for a more comprehensive security assessment of the generated protected templates, while in the *NN-learned* method it is often unclear what exactly the network learns. For the aforementioned reasons, we mainly concern the *Non-NN* BTP methods in this paper, which applies a protection algorithm to learned biometric vectors. In recent years, several *Non-NN* BTP methods [9], [10], [15], [16], [17] have been proposed, but they suffer from degradation of authentication performance due to the inevitable loss of information during transformations. In addition, these schemes are either vulnerable to security threats under the stolen-token scenario [10], [16], [17], or at risk of privacy leakage [9], [15].

To mitigate the limitations of existing schemes, we propose a template protection scheme for face and iris by combining deep hashing with cancelable distance-preserving encryption, which provides high authentication accuracy while adding high template security. A novel decision-level fusion method is designed, which uses the key generated from the iris trait to transform the face trait into a protected template. Specifically, deep hashing based architecture that can minimize the quantization loss is employed in original traits to extract compact binary codes. Then, a user-specific key can be generated by applying the reusable fuzzy extractor given the extracted iris binary code as input. Next, a cancelable distance-preserving encryption (CDPE) method is proposed to transform the face binary code into a protected template using the user-specific key, which can provide authentication accuracy equivalent to that of the unprotected system as no additional bit errors are generated. To further improve the template security, different from typical cancelable biometrics which need secure key storage, we store the digital lockers that lock the user-specific key. With this construction, the key can be unlocked from the digital lockers only if the probe iris trait is sufficiently similar to the enrollment trait. Therefore, the authentication decision of our scheme is jointly determined by iris and face, and the authentication can only be accepted if the probe iris can recover the correct key and the probe face is close to the genuine face. The contributions of this paper are summarized as follows:

- A cancelable multi-biometric template protection scheme based on deep hashing and CDPE is proposed, where deep hashing can minimize the quantization loss and CDPE ensures that no additional bit errors are generated in the protected template, thus providing reliable template protection without degrading the authentication accuracy.
- Digital lockers based reusable fuzzy extractor is constructed to avoid directly storing user-specific keys and help to verify the legitimacy of the probe iris, which enhances the security of biometric templates.
- We evaluate the performance of the proposed scheme on actual face and iris datasets. Experimental results show that our scheme achieves an equal error rate of 0.23%, and the genuine accept rate reaches 97.54% at a false accept rate of 0.1%.
- Theoretical and experimental secure analysis demonstrate that the proposed biometric template protection scheme satisfies irreversibility, revocability and unlinkability, and is resistant to decodability and similarity attacks.

Organization: The remainder of this paper is organized as follows. Section II briefly reviews related works, and Section III presents preliminaries. Section IV gives the descriptions of the system model, threat model, and design requirements. Section V is the main part where we introduce our proposed scheme in detail. In Section VI, the performance evaluation is presented, followed by the security analysis in Section VII. The discussion is given in Section VIII, and Section IX is the conclusion of our work.

II. Related Work

We review existing schemes for protecting biometric templates in this section.

A. Biometric Template Protection

BTP methods are roughly classified into three types, namely (i) biometric cryptosystems (ii) cancelable biometrics (iii) hybrid methods.

Biometric cryptosystems use cryptography to protect templates. Cimato et al. [18] presented a modular biometric cryptosystem that protects the biometric template by XORing the key and template generated in the first and second modules, respectively. Besides, many fuzzy vault or fuzzy commitment based biometric cryptosystems [19], [20], [21], [22] have been proposed for protecting biometric templates.

Cancelable biometrics use irreversible transformations for protecting templates, which is first proposed by Ratha et al. [23]. Subsequently, various transformation methods have been proposed to generate cancelable biometric templates, such as biohashing [24], non-invertible transforms [23], salting [25] and random projections [26]. A template protection approach that can against cross-matching attacks is designed by Rathgeb et al. [27]. Index-of-Max hashing technique was employed by Jin et al. [28] to generate protected biometric templates. Ghammam et al. [29] proposed a random projection-based BTP scheme, where the projection matrix is generated from the user-specific value. Recently, the Local Sampling Code technique was used

by Sadhya et al. [30] to obtain a protected template from iris traits.

A hybrid method that protects biometric templates integrates the cancelable biometrics and the biometric cryptosystem. A hybrid system that combines a Biohashing instance with a key binding method using NXOR operations was proposed by Ao and Li [31]. Yang et al. [32] presented a Biohashing based biometric cryptosystem, which is used to protect finger-vein features generated by Gabor filters. In [33], cancelable templates are first generated using discrete Fourier transform, which are then protected by the biometric cryptosystem.

B. Deep CNN Based Biometric Template Protection

CNN has been widely employed in image classification and face authentication due to its impressive learning capabilities. The privacy of CNN-extracted features has led to a surge in demand for CNN-based biometric template protection. In general, there are two categories of CNN-based template protection methods: *Non-NN* and *NN-learned*.

The *Non-NN* is a two-step approach that first uses a trained CNN to learn embeddings from biometric traits, and then uses a secure protection algorithm to convert the learned embeddings into protected templates. For example, in [34], face embeddings extracted by CNN in different face regions are first fused, then the protected template is generated using the bio-convolving encryption on the fused feature vector. Kim [11] proposed IronMask, a real-valued error correcting code-based face template protection method, which is only suitable for the angular distance metric. Rathgeb et al. [9] presented a fuzzy vault-based scheme, in which a transformation method was introduced to convert real-valued deep face embeddings into feature sets with integer values. Recently, Vedrana et al. [17] proposed a BTP method called PolyProtect that applies multivariate polynomials to transfer face embeddings into lower-dimensional representations. Dong et al. [16] put forward a cancellable face identification scheme based on their proposed deep rank hashing, which is an improvement on Linear Subspace Ranking Hashing (LSRH).

The *NN-learned* approach requires training a CNN to learn the secure transformation algorithm, and then converts biometric traits to its protected template using the trained CNN. In [7], different biometric samples of the same user are mapped to a unique code using a specially trained CNN, hence the user authentication can be performed by comparing the unique code extracted from the probe trait. Similar to [7], Jami et al. [35] applied random perturbations to the extracted feature vectors to increase the security of the original features, and then trains a CNN that converts the perturbed vectors to predefined binary codes. Since the protected templates are pre-defined, the main limitation of [7] and [35] is that the CNN needs to be retrained for each user once the protected templates have been compromised or a new user wants to register. Hence, Talreja et al. [36] presented a neural network framework which learns user-specific representations of protected templates. The CNN is first trained to learn a binary code from face images, then the Neural Network Decoder is applied for error correction. The error-corrected binary code is finally converted into the protected template using a cryptographic hash. Recently, Mai et al. [12] incorporated randomness to the training of the CNN, so that a protected template can be directly generated from the trained CNN by taking the biometric data and user-specific keys as input. Shahreza [37] presented multi-layer perception (MLP)-hash, which generates protected templates by passing the extracted features through a user-specific randomly weighted MLP and binarizing the MLP output.

C. Multi-Biometric Template Protection

In order to provide higher accuracy and resilience against spoofing attacks, many multi-biometric template protection schemes that fuse two or more biometrics have been proposed, where fusion can be performed on feature level, decision level or score level. For example, Nagar et al. [38] presented a feature-level fusion architecture which generates a secure sketch from multiple biometric templates to protect user privacy, and provided the practical implementations of the framework employing fuzzy vault [39] and fuzzy commitment. A multi-biometric method was recently proposed by Chang et al. [40], which integrating fuzzy extractor technique with their proposed bit-wise encryption to generate protected templates. However, this approach pays little attention to the renewability of the generated keys, and does not give an efficient binarization transformation method for the input data.

In addition, cancelable biometric based multi-biometric template protection methods have also been proposed. Canuto et al. [41] presented a method based on decision-level fusion, which protects speech and iris templates using cancelable transformations. A novel framework based on Bloom filters has been developed by Marta et al. [42] to ensure protection of multi-biometric templates, and Yang et al. [43] have introduced an innovative cancelable multi-biometric system that combines fingerprint and finger-vein modalities, which offers both template protection and revocability. In addition, Keshav et al. [44] designed a feature-level fusion cancelable multi-biometric system to generate revocable and non-invertible templates, in which fingerprint and iris are combined using a projection-based approach. However, these cancelable methods enhance the security of templates with a degradation of the authentication accuracy.

As for deep CNN based multi-biometric template protection, Talreja et al. [10] employed deep hashing technique [45] to generate compact binary codes from raw face and iris images, which are then transmitted into the forward error correction decoder to generate protected templates. Kim et al. [46] proposed a deep learning based multimodal cancelable biometric scheme called CSMoFN, which takes face and periocular region as input to produce a protected template. However, these two schemes need secure storage of user-specific parameters, which will suffer from stolen-key scenario and privacy leakage.

III. Preliminaries

Two important primitives employed in this paper are introduced in this section.

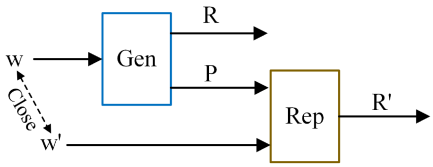


Fig. 2. Construction of fuzzy extractors.

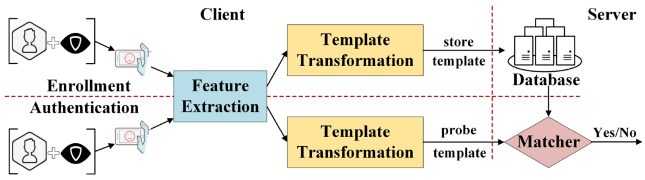


Fig. 3. System model.

A. Fuzzy Extractor

A fuzzy extractor (FE) consists of two procedures [47], a generating procedure Gen and a reproducing procedure Rep (shown in Fig. 2) with the following properties.

- Gen (w) takes $w \in \mathcal{M}$ as input, and outputs a random string $R \in \{0,1\}^n$ and a helper data $HD \in \{0,1\}^*$, i.e., Gen $(w) \rightarrow (R, HD)$. $\mathcal{M}$ represents the Hamming distance metric space.
- Rep (HD, w') on input public helper string $HD \in \{0,1\}^*$ and $w' \in \mathcal{M}$, outputs a reproduced string R' or $\perp$, i.e., Rep $(HD, w') \rightarrow R'/\perp$.

Correctness. For any Gen $(w) \rightarrow (R, HD)$ and Rep $(HD, w') \rightarrow R'$, if $dis(w, w') \leq t$, R' is equal to R, which means R can be accurately reproduced.

Security. The security requirement is that if the random source has enough entropy, then R is uniformly random.

Reusability. A reusable fuzzy extractor (rFE) retains its security and usability even if Gen is performed multiple times for the same w. We follow the definition of reusability of FE in the work of Boyen et al. [48]: For multiple samples $w, w_1, \ldots, w_\rho$ from a random source, the Gen algorithm generates corresponding $(R, HD), (R_1, HD_1), \ldots, (R_\rho, HD_\rho)$. Its security requirement is that for $i \neq j$, $HD_i \neq HD_j$, Rep$(w, HD_j) \rightarrow R_j$. And there is still pseudo-randomness among all the rest R_i.

B. Digital Locker

Digital Locker (DL) is a symmetric encryption method which is computationally secure even with weak keys [49]. We review the simple and efficient construction of digital lockers proposed by Lynn et al. [50], in which the secret key key is locked and unlocked using the value v by the following two algorithms.

- Lock(v, key) takes the value v and the secret key key as input, outputs the pair $nonce, H(nonce, v) \oplus (key||0^\gamma)$. H is a cryptographic hash function, γ represents security parameters, $nonce$ is a random number, and $||$ denotes concatenation.
- Unlock$(v', nonce, H(nonce, v) \oplus (key||0^\gamma))$ outputs correct key if $v' = v$, otherwise returns $\perp$.

With DL, it is computationally difficult to obtain any information of key given $nonce, H(nonce, v) \oplus (key||0^\gamma)$, since it is hard to guess v, and an incorrect v is easy to be recognized. The concatenation with 0^γ is utilized to ensure that Unlock can distinguish (with certainty $1 - 2^\gamma$) when the right key has been unlocked.

IV. System and Threat Model

A. System Model

Our proposed multi-biometric template protection system comprises a client and a server, which is shown in Fig. 3. The client generally refers to devices such as smartphones and laptops that are able to extract user biometric trait and have a unique identifier ID, while the server is an entity that provides online services for users. In our proposed scheme, the client stores the CNN-based feature extraction model and cancelable template transformation materials, while the server performs the matching between the enrollment and probe templates.

B. Threat Model

In a template protection system, it is critical for users to ensure the privacy of biometrics and authorized access. In a real scenario, attackers are able to eavesdrop on messages sent between the client and the server, so they may attempt to recover the raw biometric data or get linkability information from the captured messages. In addition, an attacker may want to gain authorized access from the server by providing forged biometric traits. It should be noted that the attacker is considered to be unable to capture the legitimate user's biometric traits on the client without being noticed.

C. Design Requirements

A BTP approach should satisfy four requirements according to the analysis in [51]:

1) Irreversibility: The irreversibility means that it is computationally infeasible to reverse the protected templates to obtain the raw biometric traits. This property is required to ensure biometric privacy and the security of authentication systems.

2) Revocability: The revocability represents the ability to revoke a compromised template and generate a new cancelable template. This property is required to ensure the availability of the BTP approach and avoids unauthorized access.

3) Unlinkability: The unlinkability requires that it is impossible for attackers to distinguish whether the protected templates are from the same biometric. This property is vital to prevent the leakage of user's biometric data and cross-domain matching attacks.

4) Performance: The performance of biometric-based authentication systems with template protection should not degrade significantly compared to baseline systems (authentication systems without template protection). Most of the existing

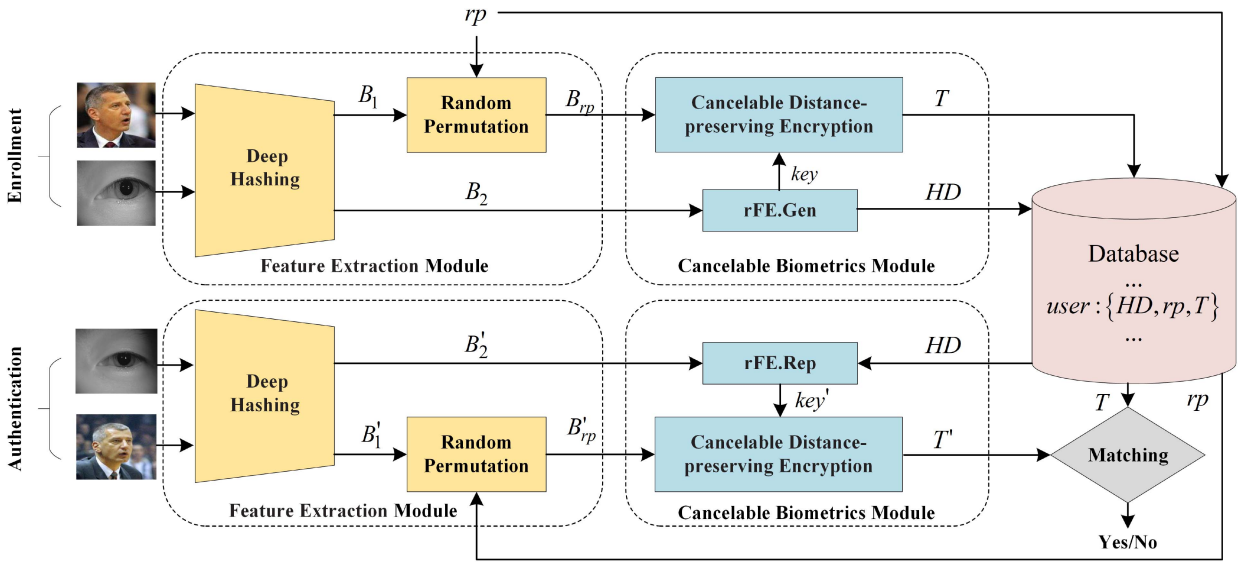


Fig. 4. Proposed cancelable multi-biometric template protection scheme.

schemes cannot satisfy this property as the non-linear nature of their transformation methods.

V. Proposed Secure Multi-Biometric System

A. System Overview

Fig. 4 demonstrates the proposed cancelable multi-biometric template protection scheme for face and iris biometrics, which mainly includes two modules: feature extraction module and cancelable biometrics module. Face and iris biometric traits are chosen in this paper for the following reasons. On the one hand, face is one of the most commonly used biometrics in our daily lives for authentication [52]. On the other hand, iris is widely regarded as the most reliable biometrics for authentication as iris remains mostly invariant throughout the lifespan of an individual [53], and the high complexity of iris textures makes iris based models resilient to false matches [54]. In addition, compared with behavioral traits (such as gait, touchscreen), face and iris are more distinguishable [55] and stable [56]. Moreover, these two traits can be easily captured simultaneously and possess complementary identity information [57], which is more convenience and user-friendly than biometrics that require additional user interaction (such as fingerprints and palmprints). Table I lists the notations used in our paper.

During enrollment, a user provides her face and iris images to the feature extraction module to extract binary codes, in which the deep hashing block consists of two CNN models (i.e., Face-CNN and Iris-CNN). The output of the feature extraction module are m-dimensional face binary code B_1 and k-dimensional iris binary code B_2. Besides, a random permutation is performed on B_1 to obtain an intermediate face binary code B_{rp} in the feature extraction module. In the cancelable biometrics module, B_2 is submitted to the rFE to extract a key and generate public helper data $HD = (mask, nonce, locker)$. It should be noted that the key is user-specific, which can be unlocked by a new iris binary code B_2' that similar to B_2. Then, a protected template T is generated given input key and B_{rp} to the CDPE method. The random permutation and the renewable user-specific keys help in achieving revocability. Once the key or random permutation vector has been compromised, a new one can be generated.

TABLE I
Table of Notations

Notation	Definition
B_1	The face binary code extracted by deep hashing
B_2	The iris binary code extracted by deep hashing
B_{rp}	The intermediate binary code after random permutation
$dis\,(\mathbf{x}, \mathbf{y})$	The Hamming distance between $\mathbf{x}$ and $\mathbf{y}$
rp	The random permutation vector
key	A user-specific key generated by rFE
k	The length of iris binary code and key
m	The length of face binary code
HD	The helper data generated by rFE
T	The protected biometric template
$H()$	A cryptographic hash function
$\|\|$	Concatenate operation
t	The number of errors that can be tolerated
τ	The decision threshold during authentication
n	The number of blocks
b	The size of each block in B_{rp}
l	The size of each block in key

During authentication, probe face and iris images presented by enrollmented user are submitted to the deep feature extractor block to generate binary codes B_1' and B_2'. A new intermediate face binary code $B_{rp}{}'$ is generated using the random permutation vector rp and B_1'. If B_2' and B_2 are similar, the secret key key can be unlocked using the HD and B_2'. Then the $B_{rp}{}'$ and key are passed through the CDPE method to obtain T'. Since the proposed multi-biometric scheme is a decision-level fusion, the authentication can only be accepted if both the probe face and iris images are similar to the enrollment images, otherwise the authentication is rejected.

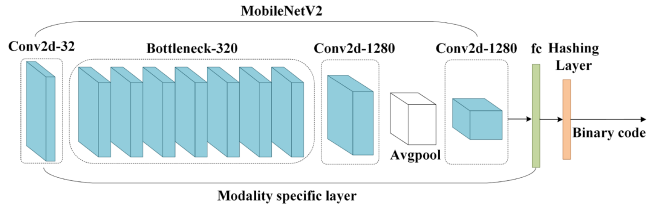


Fig. 5. Proposed architecture of deep hashing network.

B. Feature Extraction Module

The feature extraction module consists of two blocks: deep hashing block and random permutation block. The basic goal of this module is to generate two binary codes based on face and iris traits.

1) Deep Hashing: The primary goal of the deep hashing technique is non-linear feature extraction, binarization and minimizing quantization loss. The proposed architecture of the deep hashing network is shown in Fig. 5, which consists of modality specific layer and hashing representation layer.

a) Modality Specific Layer: Modality specific layer performs convolution operations on the specific biometric trait to extract features. Our work has two CNN for feature extraction, one for the face trait (Face-CNN) and another for iris trait (Iris-CNN). For each CNN, we use MobileNetV2 with an additional fully connected layer as the basic modality specific layer and train it using the real face or iris dataset, i.e., VGG-Face2 [58] for face-CNN and CASIA-Iris-Thousand for iris-CNN. The role of the additional fully connected layer in the proposed architecture is to transform the number of extracted bits into a pre-defined number. The reasons for choosing MobileNetV2 is that the number of parameters of MobileNetV2 is less than other architecture such as ResNet, and the authentication accuracy is higher using MobileNet2 for biometric authentication. Moreover, it makes the work highly reproducible when using a well-known architecture for both modalities.

b) Hashing Layer: The output feature vectors of modality specific layer are binarized in the hashing representation layer. In particular, the output of the fully connected layer produces a feature vector of real values. One feasible approach to binarize these real values is to threshold them based on the overall average. However, this approach will result in quantization loss and sub-optimal binary codes. In order to minimize quantization loss and generate compact binary codes, we add a new layer after the modality specific layer, i.e., hashing layer.

Specifically, the real-valued and continuous deep representations can be binarized into binary codes using the sign activation function $sgn(x)$ at the hashing layer. However, for all real-valued inputs, the gradient is zero due to the non-smoothness of $sgn(x)$, which makes it impossible for standard back propagation. Therefore, we employ $\mu = \tanh(\beta x)$ at the hashing layer, which is smoothed and can avoid the zero-gradient problem. $\beta > 0$ is an important scaling parameter and $\tanh(\beta x)$ can approach $sgn(x)$ infinitely as the β increase. In this way, the CNN can be trained by starting with a smaller β and making μ more non-smooth by increasing the β. Hence, there is a relationship between $sgn(x)$ and tanh

$$\lim_{\beta \to \infty} \tanh(\beta x) = sgn(x). \tag{1}$$

Therefore, as $\beta \to \infty$, the optimization problem in the CNN training will converge to the binarization problem with $sgn(x)$ as activation function. Therefore, we design an optimization method for training the network using the continuation methods described above. We apply the activation function $\tanh(\beta x)$ of $\beta = 1$ to the hashing layer, and start training until the CNN converges, then we retrain the converged network by increasing the value of β. The training process needs to be repeated many times until the hash layer can actually extract the binarized code.

2) Random Permutation: In order to ensure the unlinkability between the resulting protected templates, we use a random permutation operation to permutate the original vectors, then the permutated vector is sent to cancelable biometric module. Specifically, random permutation vector rp is applied into B_1 to randomly permutate the bits in the original face binary code to obtain B_{rp}. Different intermediate face binary codes can be generated using different permutation vectors, and the randomness of the permutation vectors ensures that linkage messages cannot be obtained through the generated templates. In addition, random permutation operation can also preserve distances well, that is, intra-user variations of biometric traits are well preserved in the permuted templates as the same rp is applied. Therefore, random permutation can provide unlinkability for protected templates without degrading authentication performance.

C. Cancelable Biometric Module

The cancelable biometric module contains two important blocks: key generation and cancelable distance-preserving encryption. The basic goal of this module is to generate a protected template by taking the intermediate face binary code B_{rp} and the extracted iris binary code B_2 as input.

1) Key Generation: A user-specific key is generated in this block by taking the iris binary code as input to the rFE. Specifically, we construct rFE based on the digital locker [50], and Algorithms 1 and 2 illustrates the specific construction steps.

During enrollment, a random secret key key of length k and helper data HD are generated by taking iris binary code B_2 as input to rFE, which is shown in Algorithm 1. Specifically, the key is locked using the AND value (v in the Algorithm 1) of the mask value $mask$ and B_2, where $mask \in \{0,1\}^k$ is used to eliminate errors between B_2 and $B_2{}'$. It is worth noting that a mask may not always have the ability to eliminate all errors between B_2 and $B_2{}'$, even if the Hamming distance between them is small. Therefore, multiple sets of $(mask_i, nonce_i, locker_i)$ $(i \in N)$ are required to guarantee that the right secret key can be unlocked with high probability from at least one of the lockers, where $locker$ represents the value of $H(nonce, v) \oplus (key||0^\gamma)$. N is the number of digital lockers, and the calculation of N is discussed in Section VI-A.1.b. The HD is stored in the database and will be used to unlock the key during the authentication phase, and the key is user-specific, which will be used to generate protected templates.

Algorithm 1: rFE Gen Algorithm.

Input: Iris binary code B_2.
Output: the helper data HD, the secret key key
1: $Client$ samples a random secret key key
2: **for** i in range(N) **do**
3: $mask_i \leftarrow \{0,1\}^k$
4: $v_i = mask_i \wedge B_2$
5: $nonce_i \leftarrow \{0,1\}^k$
6: $locker_i = H(nonce_i, v_i) \oplus (key||0^\gamma)$
7: **end for**
8: **return** key, $HD = (mask_i, nonce_i, locker_i)_{i\in N}$

Algorithm 2: rFE Rep Algorithm.

Input: $HD = (mask_i, nonce_i, locker_i)_{i\in N}$, iris binary code $B_2{}'$.
Output: the secret key key'
1: **for** i in range(N) **do**
2: $v_i' = mask_i \wedge B_2'$
3: ${key_i}' = H(nonce_i, v_i') \oplus locker_i$
4: **if** ${key_i}'[k:] == 0^\gamma$ **then**
5: $key' = {key_i}'[0:k]$
6: **else**
7: $key' = \perp$
8: **end if**
9: **end for**
10: **return** key'

During authentication, the key is unlocked by taking the probe iris binary code B_2' and HD as input. As shown in Algorithm 2, if $dis(B_2, B_2') \leq t$, the unlocking process in the Rep is implemented as $key||0^\gamma = H(nonce_i, {v_i}') \oplus locker_i$, where t is the amount of error that can be tolerated. Since key concatenated with 0^γ is encrypted into the locker in the Gen, the correct decryption of key' can be verified in the Rep. If B_2' and B_2 are not similar, the key cannot be unlocked, thus resulting in a failed authentication. Therefore, digital lockers based reusable fuzzy extractor can not only avoid directly storing keys to prevent user from privacy leaking, but also help to verify the legitimacy of the probe iris.

2) Cancelable Distance-Preserving Encryption: The cancelable distance-preserving encryption block generates a protected template by taking the user-specific key key and the intermediate face binary code B_{rp} as input.

Specifically, given the key and B_{rp}, a bit of the B_{rp} is encrypted by XORing with a temporary key generated from the key, and finally a protected template T is generated. We define this mapping of each input template B_{rp} to an output template T as cancelable distance-preserving encryption. As shown in Formula (2), although there are two cases for the mapping of each bit, the mapping of bits in each block of B_{rp} is kept consistent, that is, one of the cases is fixed. Therefore, the proposed encryption method preserves the distance of original templates in the transformed domain. In other words, the distance between the protected templates T and T' generated by the CDPE method

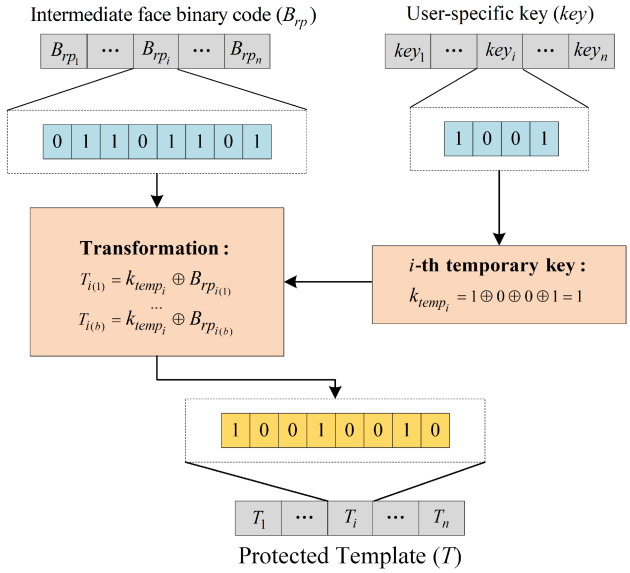


Fig. 6. Process of cancelable template generation.

is the same as that of the original biometric templates B_{rp} and ${B_{rp}}'$, i.e., $dis(T, T') = dis(B_{rp}, {B_{rp}}')$. Next, we will introduce the CDPE method in detail.

$$\begin{cases} 1_{(j)} \rightarrow 0_{(j)}, \text{ then } 0_{(j)} \rightarrow 1_{(j)} \\ 1_{(j)} \rightarrow 1_{(j)}, \text{ then } 0_{(j)} \rightarrow 0_{(j)} \end{cases}. \tag{2}$$

The proposed cancelable distance-preserving encryption method is presented in Fig. 6, which is actually a method of bit-wise encryption for biometric binary code. First of all, the m-dimensional intermediate face binary code B_{rp} and the k-dimensional key generated by rFE are split into n blocks, and the size of each block is b bits and l bits, respectively. Therefore, n should satisfy $n \times b = m$ and $n \times l = k$. We represent block sets of B_{rp} and k via $\mathcal{B}$ and $\mathcal{K}$. Thus,

$$\mathcal{B} = \left\{B_{rp_i}\right\} | i = \{1, 2, \ldots, n\}; \left|B_{rp_i}\right| = b \tag{3}$$

$$\mathcal{K} = \{key_i\} | i = \{1, 2, \ldots, n\}; |key_i| = l, \tag{4}$$

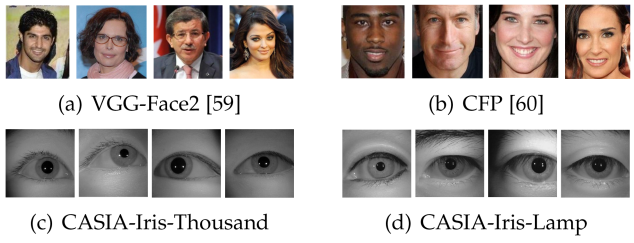
where, B_{rp_i} and key_i are individual blocks and $|.|$ denotes its size. For each block key key_i, an independent temporary key is generated via XORing each bit in key_i. Let the temporary key be represented by k_{temp}, thus,

$$k_{temp_i} = key_{i(1)} \oplus key_{i(2)} \oplus \ldots key_{i(l)}. \tag{5}$$

After the k_{temp_i} is generated for block i, the next step is to transform the i block in B_{rp} to the protected template block T_i. Specifically, the k_{temp_i} is XORed with one bit $B_{rpi(j)}$ in the ith block to generate one bit encrypted value $T_{i(j)}$

$$T_{i(j)} = k_{temp_i} \oplus B_{rpi(j)}. \tag{6}$$

Then, T_i is obtained by $T_i = T_{i(1)}||T_{i(2)}||\ldots||T_{i(b)}$. Finally, the resulting protected template T is generated by $T = T_1||T_2||\ldots||T_n$.



(a) VGG-Face2 [59] (b) CFP [60]

(c) CASIA-Iris-Thousand (d) CASIA-Iris-Lamp

Fig. 7. Example images for four datasets: (a) VGG-Face2; (b) CFP; (c) CASIA-Iris-Thousand; (d) CASIA-Iris-Lamp.

D. Similarity Score Generation

To complete authentication, the similarity score between the probe and enrollment templates is calculated and compared with the decision threshold. In the proposed scheme, we compare the similarity of blocks in the two templates one-to-one to evaluate the overall similarity between them. For two templates T and T' from the enrollment and authentication, two sets T_E and T_A of size n are formed by splitting them. Next, elements in two sets are compared one by one to calculate the number of matches, which is then divided by b to generate a normalized similarity score for the current block. When the similarity scores are calculated for all blocks, the final similarity score $Score$ can be obtained by computing the global average. We use $T_E = \{e_1, e_2, \ldots e_n\}$ and $T_P = \{p_1, p_2, \ldots p_n\}$ to represent the sets of block components for the enrollment and probe templates, respectively. Thus, the final similarity score can be computed as

$$Score = \frac{\sum_{i=1}^{n} \left(\frac{|e_j = p_j|}{b} \right)}{n} | j = \{1, \ldots, b\}, \forall e_j \in T_E, \forall p_j \in T_P. \tag{7}$$

The similarity score $Score$ computed by (7) is in $[0, 1]$. When all bits in any two blocks match each other between the enrollment and authentication templates, the $Score$ is 1. Alternatively, if there is no block match, the $Score$ is zero.

VI. Experiment and Performance Analysis

The performance evaluations of our scheme on real face and iris datasets are provided in this section. Specifically, we first introduce the experimental settings, and then report experimental results of the performance evaluation, finally provide a comparison of our method with existing CNN-based BTP schemes.

A. Experiment Setting

1) Network and Datasets: The CNN structure used for feature extraction is MobileNetV2 with a fully connected layer and a hashing layer attached, which is shown in Fig. 5. The number of nodes in the hashing layer of the Face-CNN is 1024, and that of the Iris-CNN is 128, as 128-bit key is generated in our scheme. Next, we give a brief introduction of the four datasets used in this work, and Fig. 7 displays the example images.

- **VGG–Face2** [58] dataset consists of about 3.31 million images from 9,131 subjects. These face images exhibit significant variations in terms of pose, age, lighting, and ethnicity.
- **CFP** [59] dataset consists of about 7,000 images from 500 subjects. Each subject in this dataset has 10 frontal images and 4 profile images.
- **CASIA–Iris–Thousand** dataset comprises 20,000 iris images of 1,000 subjects, each image was sampled by IKEMB -100 camera produced by IrisKing.[1] 10 left and 10 right iris images of each subject are used for fine-tuning of Iris-CNN.
- **CASIA–Iris–Lamp** dataset consists of 16,212 iris images of 411 subjects, each image was collected sampled by a hand-held iris sensor produced by OKI. Similar to CASIA-Iris-Thousand, there are 10 left iris images and 10 right iris images.

In this paper, all of the images from these four datasets are cropped to 224 × 224 before being used, and the value of each pixel in images is in [0, 255]. We randomly selected 80% images of each user from VGG-Face2 [58] and CASIA-Iris-Thousand[2] to train Face-CNN and Iris-CNN respectively, and the remaining images excluded from the training set are utilized for testing. It is worth noting that the samples in VGG-Face2 and CASIA-Iris-Thousand are sufficient for training and testing models Face-CNN and Iris-CNN.

CFP [59] and CASIA-Iris-Lamp[2] datasets are used for performance evaluation of our scheme. To construct a multimodal dataset, we combined images from CFP and CASIA-Iris-Lamp by establishing a one-to-one correspondence. Specifically, we used 10 frontal face images of 500 subjects from the CFP dataset, and the left or right iris images of each subject in CASIA-Iris-Lamp are treated as a separate class. Therefore, a multimodal dataset with 500 subjects is constructed, and there are 10 (face, iris) pairs for each subject. In our performance evaluation, one pair of each subject is used for enrollment and the remaining 9 pairs are used for genuine authentication.

2) Parameters for Proposed Scheme:

a) Parameters for Training: To ensure the accuracy of authentication, binary codes extracted by the deep hashing network should preserve the similarity of images from the same subject and the dissimilarity of images from different subjects. Therefore, the network is trained to be optimal by using a weighted maximum likelihood loss function. First, w_{ij} and s_{ij} are defined as the weight and similarity label of each training pair (x_i, x_j), respectively, where x_i, x_j are two input images. $s_{ij} = 1$ (similar) represents the images of the training pair are from the same subject, otherwise $s_{ij} = 0$ (dissimilar). Thus, the weights of similar and dissimilar pairs can be calculated by

$$w_{ij} = \begin{cases} |\mathcal{S}|/|\mathcal{S}_1|, & s_{ij} = 1 \\ |\mathcal{S}|/|\mathcal{S}_0|, & s_{ij} = 0 \end{cases}, \tag{8}$$

where $\mathcal{S} = \{s_{ij}\}$ is the set of similarity labels for all training pairs, $\mathcal{S}_1 = \{s_{ij} \in \mathcal{S} : s_{ij} = 1\}$ and $\mathcal{S}_0 = \{s_{ij} \in \mathcal{S} : s_{ij} = 0\}$ are the sets of similar and dissimilar pairs respectively, $|.|$ represents the size of set. Then, the optimization objective of

[1] [Online]. Available: http://www.irisking.com

[2] [Online]. Available: http://www.idealtest.org/

the proposed deep hashing network is to minimize intra-user variation and maximize inter-user variation

$$\min_{\Phi} \sum_{s_{ij} \in \mathcal{S}} w_{ij} \left(\log \left(1 + e^{\eta \langle \mu_i, \mu_j \rangle} \right) - \eta s_{ij} \langle \mu_i, \mu_j \rangle \right), \quad (9)$$

in which Φ represents the set of all parameters, η is the hyper-parameter, which is set to 0.1 in this paper, $\langle \cdot, \cdot \rangle$ is the inner product. Note that our network uses the smoothed activation function $\mu_i = \tanh(z_i)$. During training stage, we employ Stochastic Gradient Descent optimizer and weight decay of 10^{-5} for optimization, in which the batch size is 8. Besides, the value of the learning rate is initially set to 0.00001.

b) Parameters for Reusable Fuzzy Extractor: The reusable fuzzy extractor is implemented using digital lockers in our scheme, and multiple sets of $(mask_i, nonce_i, locker_i)$ $(i \in N)$ are needed to ensure that the right key can be unlocked with from at least one locker. To ensure the security and reusability of the rFE, the selection of the N is crucial given t, where t is the amount of error that can be tolerated. According to [60], the number of digital lockers N should satisfy

$$N \approx k^d \log \frac{2}{t}, \quad (10)$$

where d denotes a constant $d = t/\log k$ and $k = 128$ represents the length of key. It can be known that N increases exponentially with t from (10). In our work, we limit t to be below 12 bits, and the reason will be discussed in Section VI-B.1. The security parameter γ is set to 128 to avoid brute force attacks by adversaries. Thus, the value of N can be determined according to the value of t.

c) Parameters for Transformation Function: It is noticeable that the value of b determines the value of k_{temp} and affects the generation of the final protected template. Therefore, we set the block size $b = \{16, 32, 64, 128\}$ to evaluate the influences of different b on the system performance. In fact, we determine the range of b according to its impact on the overall security of our scheme, which is discussed in Section VII-E.

3) Evaluation Metrics: Equal error rate (EER) is applied to evaluate the system performance under different error tolerance t and τ. τ is the minimum acceptable similarity score between the enrollment and probe template, i.e., the decision threshold during authentication. EER represents the value of the false acceptance rate (FAR) equal to the false rejection rate (FRR), and a low EER indicates a high accuracy. FAR indicates the probability of successfully completing user authentication without the legitimate biometric images. FRR indicates the probability that the biometric images provided by legitimate users who enrollment in the database fail to be authenticated.

B. Performance Evaluation

The proposed multi-biometric scheme is a decision-level fusion, which decides the authentication result jointly by the iris and face traits. Specifically, the probe face and iris images provided by the user can be accepted require: (a) the hamming distance between the B_2' generated by the probe iris image and the B_2 is less than t; and (b) the $score$ between the T' generated by the probe face image and the stored template T is greater than the decision threshold τ. This implies that the value of FRR and FAR are determined by both t and τ. In our proposed BTP method, it is difficult to recover the key with an forged probe iris trait. Additionally, even if the key is illegally obtained by the attacker, the pseudo face images will be rejected due to the dissimilarity with the legitimate face images.

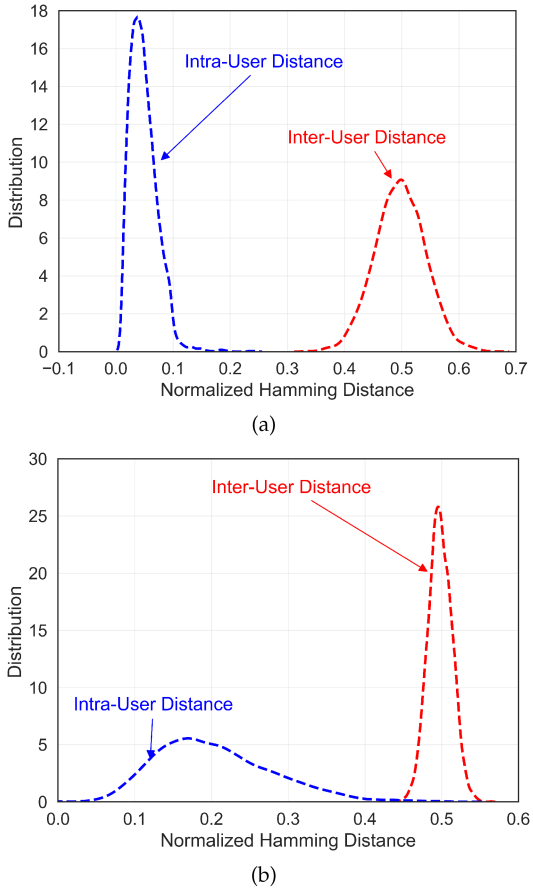
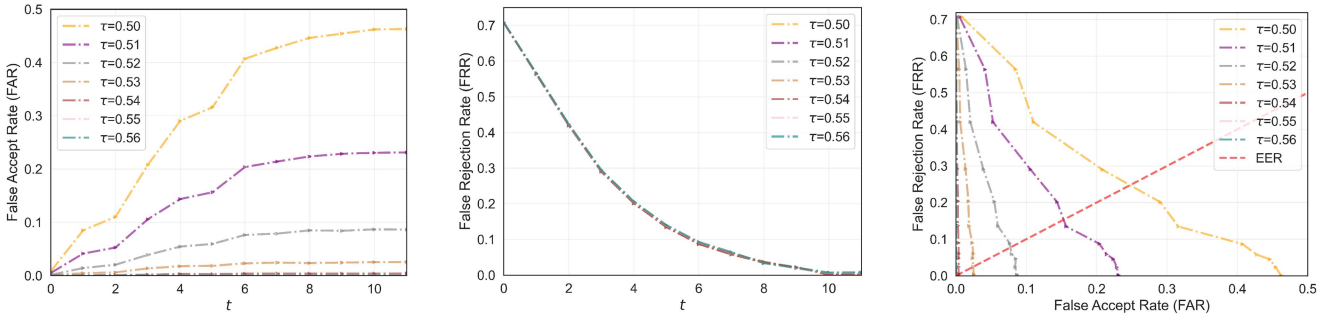


Fig. 8. Similarity evaluation based on Hamming distance: (a) intra-user and inter-user distance distribution of iris binary codes; (b) intra-user and inter-user distance distribution of protected templates.

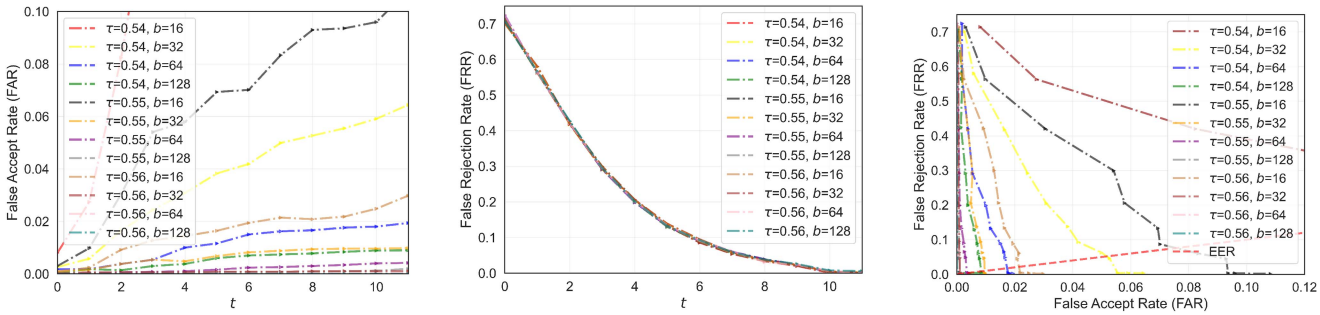
1) Similarity Evaluation: We first compare the Hamming distance of iris binary codes extracted by deep hashing between inter-user and intra-user, which reflects the possibility of an impostor recovering the user-specific secret key using forged iris traits. As the assumption described in Section IV-B, attackers cannot obtain the biometrics of legitimate users without being noticed, hence they can only attempt to break into the system by providing her own biometrics.

The distributions of hamming distance for the iris binary codes are given in Fig. 8(a), from which we can intuitively observe that the overlap between the two distributions is negligible. This indicates that iris binary codes generated by the impostor are dissimilar to that of genuine user, thus it is impossible for the impostor to recover the key using forged biometric traits. In addition, we can initially determine the error tolerance t required by iris binary codes from the distribution of the intra-user distance. Specifically, the normalized Hamming distance of the



(a) FAR (b) FRR (c) EER

Fig. 9. FAR, FRR and EER curves on different t and τ.



(a) FAR (b) FRR (c) EER

Fig. 10. Effect of different b on FAR, FRR and EER.

intra-user distance is mostly less than 0.1, and the length of the key is 128 b. Therefore, we limit t to below 12 bits in this paper.

Then we compare the Hamming distance of protected templates between inter-user and intra-user, which can reflect the possibility of an impostor passing the authentication. The distributions for the unprotected face binary codes are given in Fig. 8(b). We can see that there is a slight overlap of two distributions, so it is almost impossible for an attacker to authenticate using a spoofed biometric traits.

2) Effects of Error Tolerance and Decision Threshold: We evaluate FRR and FAR at different t and τ, which reflect the authentication accuracy of the system. We fix b to 32 and vary $t \in [0, 11]$ to analyze the authentication accuracy under different τ. As shown in Fig. 9(b), the experimental results indicate that FRR decreases as the value of t increases. This is because the error tolerance t increases, the correct key can be unlocked, so the legitimate user successfully completes the authentication. Conversely, it can be known from Fig. 9(a) that FAR increases as t increases. This is because the error tolerance t increases, some illegal users may unlock the correct key and pass the authentication within a small τ. It should be noted that both FAR and FRR should be as small as possible to ensure the authentication accuracy.

It can be analyzed that FRR and FAR are negatively correlated from the results of Fig. 9(b) and (a). In addition, both FRR and FAR are low when $t = [6, 11]$ and $\tau \in [0.52, 0.53, 0.54, 0.55, 0.56]$. Therefore, we evaluate the EER to find more appropriate t and τ to achieve a balance between FAR and FRR. The red dotted line in Fig. 9(c) denotes EER, and the points on this line mean that FRR equals FAR. The proposed scheme can reach better authentication accuracy at $t = 10$ and $\tau = \{0.54, 0.55, 0.56\}$.

3) Effects of Block Size: To find out a suitable value of τ, we further evaluate the impact of different b on system performance. For this experiment, we varied $\tau = \{0.54, 0.55, 0.56\}$, and $b = \{16, 32, 64, 128\}$. From the evaluation results of the error rate in the Fig. 10, we can know that the values of EER are small when $t = 10$, $\tau = 0.55$. In addition, as discussed in Section VII-E, the increase of b will reduce the overall security of the proposed scheme. Therefore, b is 64 in our scheme, and EER is equal to 0.0023 in this case.

4) Confusion Matrix and ROC: We evaluate the accuracy of our authentication model using confusion matrix. Specifically, iris and face data of 20 users are randomly selected to test the classification accuracy. One face image and one iris image of each user are randomly selected to generate an enrollment template, and the remaining are used as probe data for one-to-one matching with the enrollment template to get the final classification result. The above enrollment and matching process was repeated 10 times to ensure that enough classification samples were obtained, and the experimental evaluation of the confusion matrix was repeated three times to provide reliable experimental results. Fig. 11 shows that the model has a high classification accuracy, with an average of 0.9796. The reason for the

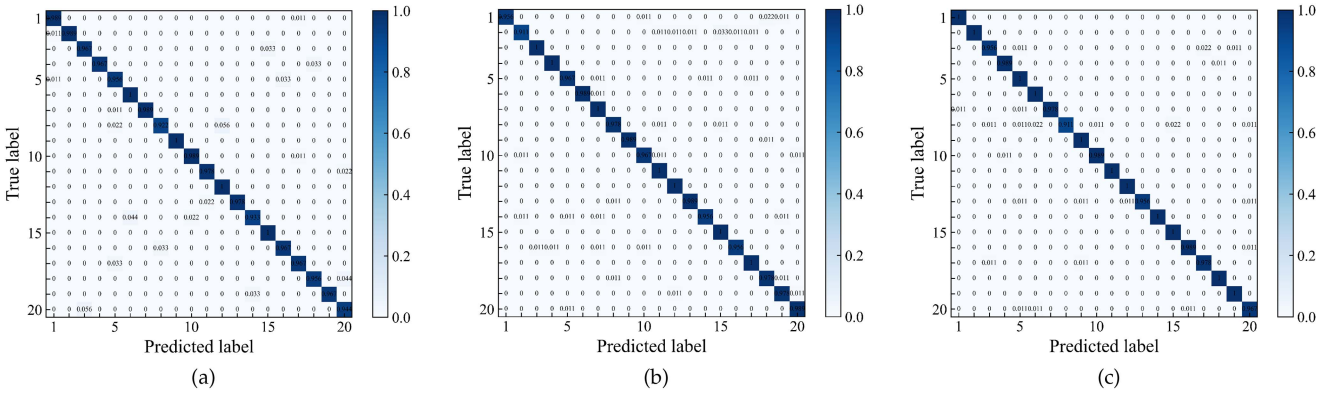


Fig. 11. Evaluation results of classification accuracy.

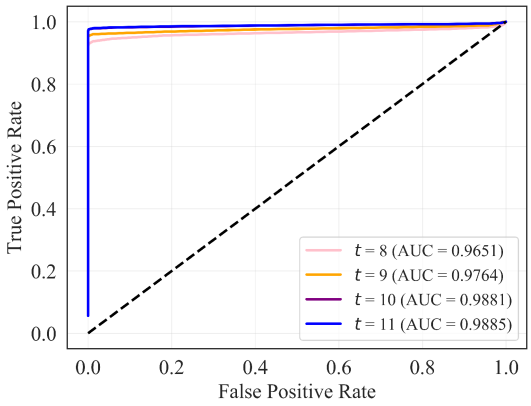


Fig. 12. ROC of different t.

misclassification is that the wrong key was recovered, so that the probe template cannot be matched to the correct enrollment template.

Furthermore, we analyze the performance of our authentication model using receiver operating characteristic curve (ROC), where τ and b are fixed to 0.55 and 64. As shown in Fig. 12, the area under ROC (AUC) exceeds 98% when t is 10 and 11, which indicates that the authentication model exhibits good classification performance and can provide high-precision user authentication.

5) Evaluation of Biometric Transposition: Moreover, we evaluate the performance of using the extracted face binary codes to generate keys to encrypt the extracted iris binary codes. Specifically, VGG-Face2 and CASIA-Iris Thousand datasets are used to train Face-CNN and Iris-CNN, where the number of nodes in the hash layer is 128 and 1024, respectively. Then, b is fixed to 64 (as discussed in Section VI-B3), and the FAR and FRR are evaluated under different τ and t using CFP and CASIA-Iris-Lamp datasets. The experimental results are shown in Fig. 13. From the evaluation results in Figs. 9 and 13, it can be seen that using iris binary codes to encrypt face binary codes has lower FRR when τ and t are the same. This is because iris is more accurate and stable than face, which helps to recover the correct key during the authentication stage, thus achieving lower FRR. Besides, the FAR is higher when $\tau \geq 0.52$ if using face binary codes to encrypt iris binary codes. Therefore, our method that first uses iris features to generate keys and then encrypts face features can achieve higher accuracy through experimental analysis.

TABLE II
COMPARISON OF OUR PROPOSED SCHEME WITH [9], [10], [12] AND [16]

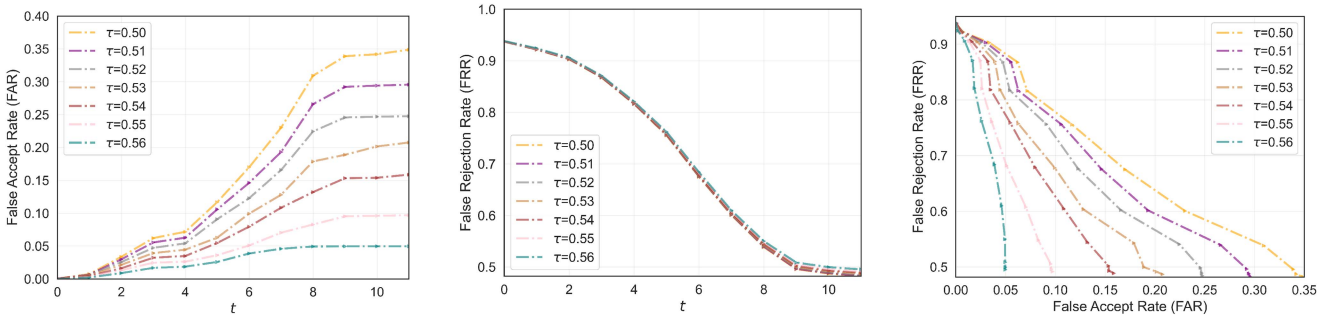
	[9]	[10]	[12]	[16]	Our scheme
Biometric	Face	Face+iris	Face	Face	Face+iris
Type	*Non-NN*	*Non-NN*	*NN-learned*	*Non-NN*	*Non-NN*
Model	ArcFace	VGG-19	ResNet-50	InceptionV2	MobileNetV2
Model size (MB)	>200	>500	>200	>1000	17.5
Security (bits)	38	104	56	-	112
GAR@0.1%FAR	95.83%	95%	85.36%	96.51%	97.54%

C. Comparison With State-of-the-Art Methods

To compare the performance of different deep hashing methods, we replace the deep hashing block in the feature extraction module of the proposed scheme with two deep hashing methods DBDH [61] and ISDH [62]. We denote the two structures as "DBDH+CDPE" and "ISDH+CDPE", and train the networks in DBDH and ISDH methods using the VGG-Face2 and CASIA-Iris-Thousand datasets described in 6.1.1. We use genuine accept rate (GAR) to evaluate the performance, where GAR denotes the probability that a genuine user is correctly accepted. Fig. 14 shows the evaluation results of GAR using CFP and CASIA-Iris-Lamp datasets, which indicates that our scheme has higher GAR than the other two systems. This is because the other two hashing techniques cannot guarantee binary strings extracted from the traits of the same user are similar, the wrong key will be locked from the probe iris binary code. Therefore, their GAR are lower than the proposed scheme.

Furthermore, we compare our scheme with recent works [9], [10], [12] and [16] that are CNN-based biometric template protection. Table II highlights the differences between our scheme and these four schemes.

Specifically, a lightweight CNN architecture MobileNetV2 is applied in the deep hashing block for feature extraction, and the security of our scheme is 112 bits (as discussed in Section



(a) FAR (b) FRR (c) FAR-FRR

Fig. 13. Evaluation results of using face features to encrypt iris features.

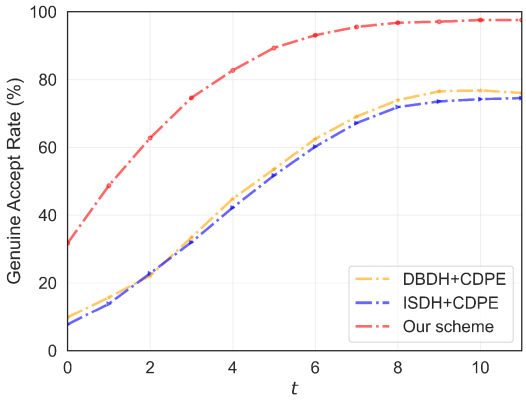


Fig. 14. GAR of different methods.

VII-E). In addition, the cancelable distance-preserving encryption method provides higher GAR and lower EER. However, CNN-based template protection method proposed in [9], [10], [12] and [16] have complex network structure, resulting in large storage overhead. Since public vaults are vulnerable to brute force attacks, the fuzzy vault-based scheme [9] has low security and cannot guarantee the unlinkability of protected templates. In [10], [16], the user-specific parameters need additional storage, which will suffer from stolen-key scenario. In addition, the *NN-learned* system [12] is limited in that it trades off authentication accuracy and the enhanced security, and it cannot protect feature embeddings generated by other CNN models.

VII. Security Analysis

The security guarantees are analyzed in this section. We first analyze the irreversibility, renewability and unlinkability. Then the security under two types of attackers and the security of the block size are analyzed.

A. Irreversibility

Irreversibility means that it is computationally difficult to reverse the enrollment templates back to the raw images. BTP methods that do not satisfy irreversibility are not available as the reversed images can be used to pass authentication. In order to obtain the raw images, a straightforward method for an attacker is brute-force, which is to guess the value of each pixel of the image. Nevertheless, this method is computationally impossible as the combinations of pixel values for each image are enormous, which is $(224 \times 224)^{256}$.

Furthermore, according to the researches [2], [4], perhaps the most feasible way to obtain reversed images is to use protected templates and helper data to learn a reconstruction model. However, in our proposed scheme, the protected template T generated by CDPE relies on both the input face image and the key. Therefore, the requirement for learning the reconstruction model is that the attacker must obtain the user-specific key first. In the key generation phase, since the attacker cannot obtain the iris biometric template B_2, they cannot unlock the correct key. To sum up, it is infeasible to get the raw biometric images by reversing the protected biometric templates.

B. Revocability

Revocability means that it is possible to revoke a compromised protected template and re-generate a new cancelable template for authentication. The revocability of our scheme is ensured by the renewable user-specific parameters, which are the key and rp. Specifically, since a reusable fuzzy extractor is constructed in the cancelable biometrics module, multiple different keys could be obtained from the same user's biometric traits, and the random permutation vector can be changed each time a new enrollment is performed. Therefore, our scheme satisfies the revocability property.

We can also experimentally verify that the generated templates are revocable. We use *pseudo-impostor* to denote the attacker who obtains a compromised biometric template of a legitimate user and attempts to obtain legal authentication. To simulate this attack, 100 protected templates were generated from the same face image by utilizing 100 different keys generated by the same iris image and 100 re-generated random permutation vectors. Subsequently, the first protected template was matched with the other 99 templates one by one to generate the *pseudo-impostor* similarity scores. The above experimental process was repeated on all 500 subjects used for testing (as described in Section VI-B). The revocability of the proposed scheme can be verified experimentally if (1) there is a clear

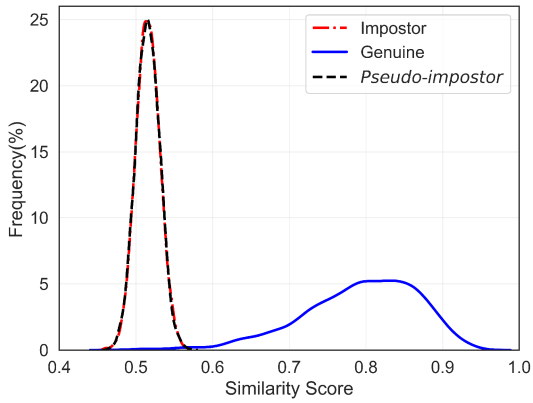


Fig. 15. Revocability analysis of the proposed scheme.

overlap between the imposter and *pseudo-impostor* distributions, and (2) there is a clear separation between the genuine and *pseudo-impostor* distributions. We completed the above experiments with optimal settings ($b = 64, t = 10, \tau = 0.55$), and Fig. 15 shows the similarity score distributions of the three cases. It can be seen from Fig. 15 that the distributions of similarity score between the *pseudo-impostor* and the impostor have a large overlap, while the *pseudo-impostor* and the genuine are clearly separated. Hence, we experimentally confirmed that the scheme satisfies revocability.

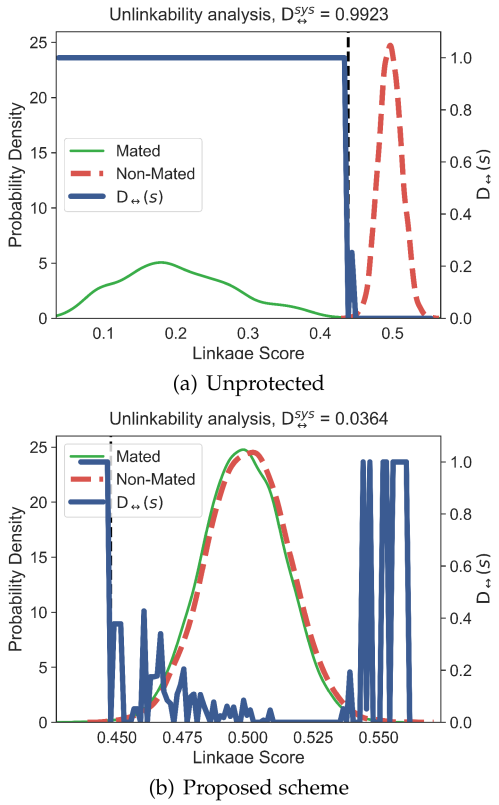


Fig. 16. Unlinkability analysis for the templates without/with protection.

C. Unlinkability

Unlinkability indicates that an attacker cannot obtain linkage information from multiple protected templates generated by the same biometric trait, i.e., the generated templates are indistinguishable.

1) Unlinkability of Reusable Fuzzy Extractor: Although the HD generated during the key generation phase is publicly stored on the server, our design meets the unlinkability. This is because the generated helper data are pseudo-random, the rFE remains secure even if attackers get the helper data generated from the same iris trait multiple times. Specifically, since the generated digital lockers are indistinguishable from the attackers under random oracle [50] and the values of $mask$ are randomly generated, the $(mask, nonce, locker)$ pairs are independent of each other and will not reveal information of the raw biometric traits. Hence, the combination of the helper data with the protected templates of the user cannot obtain linkage information.

2) Unlinkability Analysis of Template Protection Scheme: ISO/IEC International Standard 24745 [6] states that different transformation templates generated from the same biometric traits should not be linked. We evaluated the unlinkability of protected templates generated by our scheme using the protocol defined in [63]. Specifically, mated (H_m) and non-mated (H_{nm}) samples distributions are employed to analyze the unlinkability [63]. Mated samples denote protected templates generated from the biometrics of the same subject, while non-mated samples are generated from the biometrics of different subjects. There must be a significant overlap between the distributions of link scores computed from paired and unpaired samples. For a system that meets unlinkability, there must be a significant overlap between the distributions of linkage scores computed from mated and non-mated samples [63].

Two measures of unlinkability are defined based on the above two distributions of linkage score: i) Local measure $D_{\leftrightarrow}(s)$, which evaluates the unlinkability between protected templates based on the likelihood ratio between the distributions of the linkage score. The value of $D_{\leftrightarrow}(s)$ is in $[0, 1]$, $D_{\leftrightarrow}(s) = 0$ indicates that the transformation templates at the current score s are completely unlinkable, conversely, the templates are completely linkable when $D_{\leftrightarrow}(s) = 1$. ii) Global measure $D^{sys}_{\leftrightarrow}$, which evaluates the unlinkability independently of the linkage score and depends on the comparison of two or more systems. Similarly, the value of $D^{sys}_{\leftrightarrow}$ is in $[0, 1]$, $D^{sys}_{\leftrightarrow} = 0$ means complete unlinkability and $D^{sys}_{\leftrightarrow} = 1$ denotes the complete linkability for the transformed templates. For $D_{\leftrightarrow}(s)$ and $D^{sys}_{\leftrightarrow}$, the values between 0 and 1 represent the degree of unlinkability.

In this scenario, 10 protected templates were generated for each subject by using different face and iris images, and 10 unprotected templates from different face images were generated by using deep hashing. Then, different templates of the same subject are used to represent mated samples, and templates of different subjects are used to represent non-mated samples. We use the Hamming distance to calculate the linkage score, and the distribution of linkage scores for mated samples, non-mated samples and $D_{\leftrightarrow}(s)$ is shown in Fig. 16. The 'Unprotected' (Fig. 16(a)) demonstrates the evaluation results of linkage scores by using templates generated without random permutation and

CDPE. It is observed from the Fig. 16(a) that the linkability of unprotected templates is very high, as the $D_{\leftrightarrow}^{sys}$ on test datasets is 0.9923. Conversely, the global linkability measure $D_{\leftrightarrow}^{sys}$ for the protected templates is low, which is 0.0364. Therefore, the proposed BTP scheme meets the unlinkability.

D. Security of the System Under Attacks

Furthermore, decodability attacks [64] and similarity attacks [65] may break the unlinkability or irreversibility of protected templates by reducing the attack complexity. We discuss the security of the proposed scheme under these two types of attacks in this section.

1) Decodability Attack via Helper Data: As described in Section VII-C, the helper data HD_i and HD_j generated by reusable fuzzy extractor using the same biometric meets the unlinkability, an attacker can not determine whether HD_i and HD_j come from the same subject. Further, the elements in the set HD_i are uniform distribution, the attacker is impossible to get useful information from it as he can not distinguish which piece of data can recover the key. Consider a scenario where an attacker obtains a secret key, and the attacker attempts to use the key and HD to obtain the iris traits of the legitimate user. However, she can only get the $HMAC(nonce, v)$, which is difficult to recover the original iris images.

2) Similarity Attack: As described in Section VI-A.2.a, Face-CNN and Iris-CNN models are trained with the optimization objective of minimizing variations of intra-user and maximizing variations of inter-user, and the similarity evaluation in Section VI-B.1 shows that the distribution of Hamming distances between intra-user slightly overlaps with that between inter-user. Therefore, it is difficult for an attacker to attempt to perform a similarity attack using spoofed images.

In addition, similarity attacks can be done by estimating two images x' (face) and y' (iris) given the protected templates of a legitimate user, such that the $score$ between the template generated by the estimated images and the enrollment template is less than the decision threshold. The protected template T generated at enrollment depends on both the rp and the key, in which the key is locked into the HD. The key can only be unlocked correctly from HD when y' is sufficiently similar to the enrollment image. There is no advantage in generating a template T_s whose similarity score to T is less than the decision threshold if the key cannot be correctly unlocked. Therefore, the similarity attack can break the security of our scheme only if the images x' and y' are sufficiently similar to the enrollment images x and y. This is equivalent to brute-forcing the enrollment images, which is computationally complex due to the large number of combinations of pixels in each image. To sum up, the proposed scheme can resist the similarity attack.

E. Security Analysis of the Block Size

Now we theoretically analyze the impact of block size on scheme security. In our proposed scheme, the block size b affects the size l and number n of block keys, that is, $n = m/b$ and $l = (kb)/m$. In fact, as discussed in Section VII-A, the irreversibility of protected templates relies on the user-specific secret key.

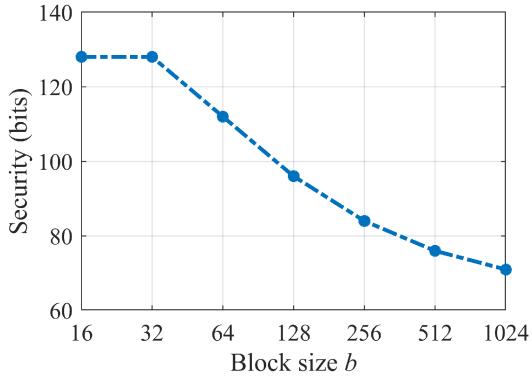


Fig. 17. Security for different value of b.

Therefore, we analyze the effect of block size on the security provided by the key.

Specifically, an attacker can perform a guessing attack by guessing n l-bit block key one by one. As described in Section V-C.2, each l-bit block key is XORed bit by bit to generate a temporary key in the CDPE method. Therefore, an attacker can first guess the value of the temporary key, which is 0 or 1. Then, every time the attacker guesses a bit in the block key, there will be one bit corresponding to the guessed bit. This is because of the property of the XOR operation, i.e., the final calculation result must be the guessed temporary key. Based on the above discussion, the attacker only needs to guess $l/2$ bits of the block key, then arrange the corresponding bits of the $l/2$ bits, and the average number of arrangement is $l/2$. Hence, the complexity of guessing all n block keys can be calculated as

$$\mathcal{C}(n,b) = \left(2 * 2^{\frac{(kb)/m}{2}} * \frac{(kb)/m}{2}\right)^n. \tag{11}$$

Where $k = 128$ and $m = 1024$ are the lengths of the iris and face binary codes, respectively. From the (11), we can obtain the security of our scheme, which is denoted as $\log_2 \mathcal{C}(n,b)$.

As shown is Fig. 17, the security decreases as b increases, which reaches a minimum of 71 when $b = 1024$. In order to choose an appropriate b for our scheme, we take both the security and authentication accuracy into account. On the one hand, a smaller b will provide higher security for the proposed scheme. On the other hand, the EER decreases with the increase of b from Fig. 10. Therefore, $b = 64$ is chosen in our scheme, where 112 bits of security is provided and EER is equal to 0.0023.

VIII. Discussion

In the proposed cancelable multi-biometric template protection scheme, deep hashing technique is employed to extract compact binary vectors from raw face and iris data, which minimizes quantization loss. Different from [12], the enrollment of new users in our scheme does not need to retrain the Face-CNN and Iris-CNN models, which increases usability. The use of the lightweight network structure MobileNetV2 with a special optimization objective provides lower storage overhead and higher accuracy than the CNN-based biometric template protection schemes [9], [10], [12], [16]. Besides, the user-specific

parameters used for template encryption is not directly stored in the database like [10] and [16], but is recovered by using probe iris images and helper data, which can prevent user privacy from leaking and avoid stolen-token scenario. Moreover, the proposed CDPE method transforms the face trait into a protected template without additional bit errors, thus providing authentication accuracy equivalent to that of the unprotected system and reliable template protection.

Limitations: Although the proposed scheme achieves multi-biometric template protection, it is only suitable for iris and face traits, which lacks generality. For other biometrics, it is necessary to train the corresponding feature extraction models based on the designed deep hashing architecture and datasets, then the important parameters τ, b, t can be determined for high accuracy and secure authentication. In addition, like other multi-biometric template protection schemes [10], [40], the proposed work lacks of flexibility due to the requirement of submitting two traits simultaneously. Users sometimes fail to provide the required traits in some extreme environments, such as dark nights.

IX. Conclusion

In this paper, we proposed a cancelable multi-biometric template protection scheme by combining the deep hashing with cancelable distance-preserving encryption. First, the deep hashing network is used to extract compact binary codes from face and iris images. Second, the iris binary code is submitted to the rFE to generate a user-specific key, and the face binary code is permuted using a random permutation vector. In order to enhance the security, the user-specific key is locked in digital lockers and can only be unlocked if the probe and enrollment iris images are similar enough. Third, a protected template is generated by taking the key and the permuted face binary code as input to CDPE. Experimental results on practical face and iris datasets show that our method can achieve an EER of 0.23%. In addition, theoretical and experimental analysis indicate that our scheme satisfies irreversibility, revocability and unlinkability, and can resist decodability and similarity attacks.

In our future work, we would like to consider employing cross-modal learning to enhance the flexibility of our scheme, as the goal of cross-modal learning is to develop a model that can efficiently extract meaningful correlation information from multiple modalities [66], [67], thus enabling flexible authentication even if one of the modalities is missing.

References

[1] Y. C. Feng, M.-H. Lim, and P. C. Yuen, "Masquerade attack on transform-based binary-template protection based on perceptron learning," *Pattern Recognit.*, vol. 47, no. 9, pp. 3019–3033, 2014.

[2] G. Mai, K. Cao, P. C. Yuen, and A. K. Jain, "On the reconstruction of face images from deep face templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 5, pp. 1188–1202, May 2019.

[3] K. Cao and A. K. Jain, "Learning fingerprint reconstruction: From minutiae to image," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 104–117, Jan. 2015.

[4] J. Galbally, A. Ross, M. Gomez-Barrero, J. Fierrez, and J. Ortega-Garcia, "Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms," *Comput. Vis. Image Understanding*, vol. 117, no. 10, pp. 1512–1525, 2013.

[5] P. Regulation, "Regulation (EU) 2016/679 of the European parliament and of the council," *Regulation*, vol. 679, pp. 1–88, 2016.

[6] ISO/IEC, JTC 1/SC 27, "Information technology–security techniques–biometric information protection," Jun. 2011.

[7] A. Kumar Jindal, S. Chalamala, and S. Kumar Jami, "Face template protection using deep convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 462–470.

[8] V. K. Hahn and S. Marcel, "Biometric template protection for neural-network-based face recognition systems: A survey of methods and evaluation techniques," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 639–666, 2022.

[9] C. Rathgeb, J. Merkle, J. Scholz, B. Tams, and V. Nesterowicz, "Deep face fuzzy vault: Implementation and performance," *Comput. Secur.*, vol. 113, 2022, Art. no. 102539.

[10] V. Talreja, M. C. Valenti, and N. M. Nasrabadi, "Deep hashing for secure multimodal biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1306–1321, 2020.

[11] S. Kim, Y. Jeong, J. Kim, J. Kim, H. T. Lee, and J. H. Seo, "IronMask: Modular architecture for protecting deep face template," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 16125–16134.

[12] G. Mai, K. Cao, X. Lan, and P. C. Yuen, "SecureFace: Face template protection," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 262–277, 2020.

[13] J. R. Pinto, M. V. Correia, and J. S. Cardoso, "Secure triplet loss: Achieving cancelability and non-linkability in end-to-end deep biometrics," *IEEE Trans. Biom., Behav., Ident. Sci.*, vol. 3, no. 2, pp. 180–189, Apr. 2021.

[14] H. Lee, C. Y. Low, and A. B. J. Teoh, "SoftmaxOut transformation-permutation network for facial template protection," in *Proc. 25th Int. Conf. Pattern Recognit.*, 2021, pp. 7558–7565.

[15] D. D. Mohan, N. Sankaran, S. Tulyakov, S. Setlur, and V. Govindaraju, "Significant feature based representation for template protection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 2389–2396.

[16] X. Dong, S. Cho, Y. Kim, S. Kim, and A. B. J. Teoh, "Deep rank hashing network for cancellable face identification," *Pattern Recognit.*, vol. 131, 2022, Art. no. 108886.

[17] V. K. Hahn and S. Marcel, "Towards protecting face embeddings in mobile face verification scenarios," *IEEE Trans. Biom., Behav., Ident. Sci.*, vol. 4, no. 1, pp. 117–134, Jan. 2022.

[18] S. Cimato, M. Gamassi, V. Piuri, R. Sassi, and F. Scotti, "Privacy-aware biometrics: Design and implementation of a multimodal verification system," in *Proc. Annu. Comput. Secur. Appl. Conf.*, 2008, pp. 130–139.

[19] T. K. Dang, Q. C. Truong, T. T. B. Le, and H. Truong, "Cancellable fuzzy vault with periodic transformation for biometric template protection," *IET Biometrics*, vol. 5, no. 3, pp. 229–235, 2016.

[20] V. S. Baghel, S. Prakash, and I. Agrawal, "An enhanced fuzzy vault to secure the fingerprint templates," *Multimedia Tools Appl.*, vol. 80, pp. 33055–33073, 2021.

[21] T. A. T. Nguyen, T. K. Dang, and D. T. Nguyen, "A new biometric template protection using random orthonormal projection and fuzzy commitment," in *Proc. 13th Int. Conf. Ubiquitous Inf. Manage. Commun.*, 2019, pp. 723–733.

[22] S. Shukla and S. J. Patel, "Securing fingerprint templates by enhanced minutiae-based encoding scheme in fuzzy commitment," *IET Inf. Secur.*, vol. 15, no. 3, pp. 256–266, 2021.

[23] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, "Generating cancelable fingerprint templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 561–572, Apr. 2007.

[24] A. Kong, K.-H. Cheung, D. Zhang, M. Kamel, and J. You, "An analysis of biohashing and its variants," *Pattern Recognit.*, vol. 39, no. 7, pp. 1359–1368, 2006.

[25] J. Zuo, N. K. Ratha, and J. H. Connell, "Cancelable iris biometric," in *Proc. 19th Int. Conf. Pattern Recognit.*, 2008, pp. 1–4.

[26] A. B. Teoh, Y. W. Kuan, and S. Lee, "Cancellable biometrics and annotations on biohash," *Pattern Recognit.*, vol. 41, no. 6, pp. 2034–2044, 2008.

[27] J. Bringer, C. Morel, and C. Rathgeb, "Security analysis and improvement of some biometric protected templates based on bloom filters," *Image Vis. Comput.*, vol. 58, pp. 239–253, 2017.

[28] Z. Jin, J. Y. Hwang, Y.-L. Lai, S. Kim, and A. B. J. Teoh, "Ranking-based locality sensitive hashing-enabled cancelable biometrics: Index-of-max hashing," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 2, pp. 393–407, Feb. 2018.

[29] L. Ghammam, M. Barbier, and C. Rosenberger, "Enhancing the security of transformation based biometric template protection schemes," in *Proc. Int. Conf. Cyberworlds*, 2018, pp. 316–323.

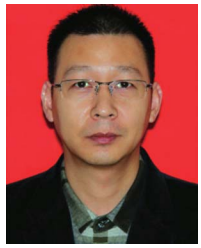
[30] D. Sadhya and B. Raman, "Generation of cancelable iris templates via randomized bit sampling," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 11, pp. 2972–2986, Nov. 2019.
[31] M. Ao and S. Z. Li, "Near infrared face based biometric key binding," in *Proc. Int. Conf. Biometrics*, 2009, pp. 376–385.
[32] W. Yang et al., "Securing mobile healthcare data: A smart card based cancelable finger-vein bio-cryptosystem," *IEEE Access*, vol. 6, pp. 36939–36947, 2018.
[33] B. Alam, Z. Jin, W.-S. Yap, and B.-M. Goi, "An alignment-free cancelable fingerprint template for bio-cryptosystems," *J. Netw. Comput. Appl.*, vol. 115, pp. 20–32, 2018.
[34] E. Abdellatef et al., "Cancelable multi-biometric recognition system based on deep learning," *Vis. Comput.*, vol. 36, no. 6, pp. 1097–1109, 2020.
[35] S. K. Jami, S. R. Chalamala, and A. K. Jindal, "Biometric template protection through adversarial learning," in *Proc. IEEE Int. Conf. Consum. Electron.*, 2019, pp. 1–6.
[36] V. Talreja, M. C. Valenti, and N. M. Nasrabadi, "Zero-shot deep hashing and neural network based error correction for face template protection," in *Proc. IEEE 10th Int. Conf. Biometrics Theory Appl. Syst.*, 2019, pp. 1–10.
[37] H. O. Shahreza, V. K. Hahn, and S. Marcel, "MLP-hash: Protecting face templates via hashing of randomized multi-layer perceptron," in *Proc. IEEE 31st Eur. Signal Process. Conf.*, 2023, pp. 605–609.
[38] A. Nagar, K. Nandakumar, and A. K. Jain, "Multibiometric cryptosystems based on feature-level fusion," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 255–268, Feb. 2012.
[39] A. Juels and M. Sudan, "A fuzzy vault scheme," *Designs, Codes Cryptogr.*, vol. 38, no. 2, pp. 237–257, 2006.
[40] D. Chang, S. Garg, M. Hasan, and S. Mishra, "Cancelable multi-biometric approach using fuzzy extractor and novel bit-wise encryption," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3152–3167, 2020.
[41] A. M. Canuto, F. Pintro, and J. C. Xavier-Junior, "Investigating fusion approaches in multi-biometric cancellable recognition," *Expert Syst. Appl.*, vol. 40, no. 6, pp. 1971–1980, 2013.
[42] M. Gomez-Barrero, C. Rathgeb, G. Li, R. Ramachandra, J. Galbally, and C. Busch, "Multi-biometric template protection based on bloom filters," *Inf. Fusion*, vol. 42, pp. 37–50, 2018.
[43] W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, "A fingerprint and finger-vein based cancelable multi-biometric system," *Pattern Recognit.*, vol. 78, pp. 242–251, 2018.
[44] K. Gupta, G. S. Walia, and K. Sharma, "Novel approach for multimodal feature fusion to generate cancelable biometric," *Vis. Comput.*, vol. 37, pp. 1401–1413, 2021.
[45] Z. Cao, M. Long, J. Wang, and P. S. Yu, "HashNet: Deep learning to hash by continuation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 5608–5617.
[46] J. Kim, Y. G. Jung, and A. B. J. Teoh, "Multimodal biometric template protection based on a cancelable SoftmaxOut fusion network," *Appl. Sci.*, vol. 12, no. 4, pp. 1–20, 2022, doi: 10.3390/app12042023.
[47] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *Proc. Int. Conf. Theory Appl. Cryptographic Techn.*, 2004, pp. 523–540.
[48] X. Boyen, "Reusable cryptographic fuzzy extractors," in *Proc. 11th ACM Conf. Comput. Commun. Secur.*, 2004, pp. 82–91.
[49] R. Canetti, Y. T. Kalai, M. Varia, and D. Wichs, "On symmetric encryption and point obfuscation," in *Proc. Theory Cryptogr. Conf.*, 2010, pp. 52–71.
[50] B. Lynn, M. Prabhakaran, and A. Sahai, "Positive results and techniques for obfuscation," in *Proc. Int. Conf. Theory Appl. Cryptographic Techn.*, 2004, pp. 20–39.
[51] N. Kumar, "Cancelable biometrics: A comprehensive survey," *Artif. Intell. Rev.*, vol. 53, no. 5, pp. 3403–3446, 2020.
[52] A. Kumar, S. Jain, and M. Kumar, "Face and gait biometrics authentication system based on simplified deep neural networks," *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 1005–1014, 2023.
[53] J. Daugman, "How iris recognition works," in *Proc. Essential Guide Image Process.*, 2009, pp. 715–739.
[54] J. Daugman, "Information theory and the iriscode," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 400–409, Feb. 2016.
[55] M. De Marsico, C. Galdi, M. Nappi, and D. Riccio, "FIRME: Face and iris recognition for mobile engagement," *Image Vis. Comput.*, vol. 32, no. 12, pp. 1161–1172, 2014.
[56] M. A. El-Bendary, H. Kasban, A. Haggag, and M. El-Tokhy, "Investigating of nodes and personal authentications utilizing multimodal biometrics for medical application of WBANs security," *Multimedia Tools Appl.*, vol. 79, pp. 24507–24535, 2020.
[57] Q. Zhang, H. Li, M. Zhang, Z. He, Z. Sun, and T. Tan, "Fusion of face and iris biometrics on mobile devices using near-infrared images," in *Proc. 10th Chin. Conf. Biometric Recognit.*, Tianjin, China, 2015, pp. 569–578.
[58] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in *Proc. IEEE 13th Int. Conf. Autom. Face Gesture Recognit.*, 2018, pp. 67–74.
[59] S. Sengupta, J.-C. Chen, C. Castillo, V. M. Patel, R. Chellappa, and D. W. Jacobs, "Frontal to profile face verification in the wild," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2016, pp. 1–9.
[60] R. Canetti, B. Fuller, O. Paneth, L. Reyzin, and A. Smith, "Reusable fuzzy extractors for low-entropy distributions," in *Proc. Annu. Int. Conf. Theory Appl. Cryptographic Techn.*, 2016, pp. 117–146.
[61] X. Zheng, Y. Zhang, and X. Lu, "Deep balanced discrete hashing for image retrieval," *Neurocomputing*, vol. 403, pp. 224–236, 2020.
[62] Z. Zhang, Q. Zou, Q. Wang, Y. Lin, and Q. Li, "Instance similarity deep hashing for multi-label image retrieval," 2018, *arXiv:1803.02987*.
[63] M. Gomez-Barrero, J. Galbally, C. Rathgeb, and C. Busch, "General framework to evaluate unlinkability in biometric template protection systems," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1406–1420, Jun. 2018.
[64] B. Tams, "Decodability attack against the fuzzy commitment scheme with public feature transforms," 2014, *arXiv:1406.1154*.
[65] Y. Chen, Y. Wo, R. Xie, C. Wu, and G. Han, "Deep secure quantization: On secure biometric hashing against similarity-based attacks," *Signal Process.*, vol. 154, pp. 314–323, 2019.
[66] A. Nagrani, S. Albanie, and A. Zisserman, "Seeing voices and hearing faces: Cross-modal biometric matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8427–8436.
[67] S. Horiguchi, N. Kanda, and K. Nagamatsu, "Face-voice matching using cross-modal embeddings," in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 1011–1019.

Guichuan Zhao received the bachelor of engineering degree from the School of Computer Science and Technology, Harbin Institute of Technology, in 2018. He is currently working toward the postgraduate degree with the School of Cyber Engineering, Xidian University, China. His research interests include security in body area networks, wireless sensor networks, biometric authentication.

Qi Jiang received the BS degree in computer science from Shaanxi Normal University, in 2005, and the PhD degree in computer science from Xidian University, in 2011. He is now a professor with the School of Cyber Engineering, Xidian University. His research interests include security protocols and wireless network security, cloud security, etc.

Ding Wang received the PhD degree in information security from Peking University, in 2017, and was supported by the "Boya Postdoctoral Fellowship" in Peking University from 2017 to 2019. Currently, he is a full professor with Nankai University. As the first author (or corresponding author), he has published more than 60 papers with venues like IEEE S&P, ACM CCS, NDSS, Usenix Security, *IEEE Transactions on Dependable and Secure Computing* and *IEEE Transactions on Information Forensics and Security*. His research has been reported by more than 200 medias like Daily Mail, Forbes, *IEEE Spectrum* and *Communications of the ACM*, appeared in the Elsevier 2017 "Article Selection Celebrating Computer Science Research in China", and resulted in the revision of the authentication guideline NIST SP800-63-2. He has been involved in the community as a PC Chair/TPC member for more than 70 international conferences such as ACM CCS 2022, NDSS 2023, PETS 2022/2023, ACSAC 2020–2022, ACM AsiaCCS 2021/2022, ICICS 2018–2022, INSCRYPT 2018–2022. He has received the "ACM China Outstanding Doctoral Dissertation Award", the Best Paper Award with INSCRYPT 2018, and the First Prize of Natural Science Award of Ministry of Education. His research interests focus on passwords, authentication and provable security.

Xindi Ma received the BS degree from the School of Computer Science and Technology, Xidian University, in 2013, and the PhD degree in computer science from Xidian University, in 2018. He is now a postdoctor with the School of Cyber Engineering, Xidian University. His current research interests include privacy computing, recommender system and machine learning with focus on security and privacy issues.



Xinghua Li (Member, IEEE) received the MS and PhD degrees in computer science from Xidian University, in 2004 and 2007, respectively. He is currently a professor with the School of Cyber Engineering, Xidian University, China. His research interests include wireless networks security, and protocol formal methodology.