A Secure and Privacy-Preserving Biometric Template Protection Technique

Anonymous IJCB 2023 submission

Abstract

Vein patterns are an important source of biometric information due to their uniqueness and permanence. As with any other biometric modality, the security and privacy of vascular biometric templates are of utmost importance to prevent unauthorized access and identity theft. We propose a privacy-preserving method using SVM classifiers and having them replace the stored template to enhance the security and privacy of biometric data with minimal impact on accuracy. As a bonus, the proposed system generates stable, non-invertible keys enabling a host of cryptographic functionalities. Cosine similarity matching of extracted features showed an area under the curve (AUC) of 99.97% under the receiver operating characteristic (ROC) curve and an EER of 0.6% on the test set. Our proposed method showed an AUC of 99.9% under the ROC curve, indicating excellent discrimination ability. In addition, the equal error rate is 0.673% on the test set, suggesting that our method has achieved a good balance between security and usability, with a low rate of both false positives and false negatives. These results suggest that the privacy and security of biometric templates can be improved without compromising system performance.

1. Introduction

Biometric authentication systems utilize an individual's unique physical or behavioral characteristics to establish their identity [12]. Unlike traditional authentication methods, such as passwords or PINs, biometric authentication does not require users to memorize or carry any authentication tokens. Instead, the system compares the biometric traits of an individual with those stored in a database to determine their identity.

Biometric traits can be classified into two categories: physical and behavioral. Physical traits include fingerprints, iris patterns, facial recognition, hand geometry, and vein patterns. Behavioral traits include voice recognition, keystroke or touchscreen dynamics, gait, and signature recognition.

Biometric authentication systems provide enhanced convenience, accuracy, and security, as they rely on unique physiological traits that are difficult for imposters to replicate. However, concerns exist regarding the privacy and security of biometric data, as it is a sensitive asset that requires strict adherence to privacy and security guidelines to prevent unauthorized access, misuse, or theft. Furthermore, biometric authentication systems are vulnerable to various attacks such as presentation attacks, where an imposter attempts to mimic the biometric traits of a legitimate user. To address these concerns, it is crucial to implement robust security measures to protect biometric data, such as encryption, secure storage, and access controls. Storing biometric templates, especially in their original form, creates untenable security and privacy issues.

Vein biometrics, also known as vascular biometrics, is an important biometric modality due to its inherent accuracy and stability. As an internal biometric, vein patterns are highly complex and unique to individuals, making them extremely difficult to forge or replicate. Subdermal vascular patterns can be captured using near-infrared (NIR) light. When near-infrared (NIR) light is directed at a part of the body, the deoxyhemoglobin in the blood present in the veins absorbs the light, resulting in a distinct pattern of dark lines that can be captured using readily available NIR cameras.

Vein biometric recognition offers several advantages over other biometric technologies such as [9]:

1. *Uniqueness*: Extensive research has demonstrated that the diversity of vein patterns is comparable to that of iris patterns, which makes finger-vein recognition an exceptionally secure method for biometric authentication. Even identical twins express different vascular patterns.
2. *Permanence*: Most vascular patterns, such as finger-vein patterns, remain constant throughout adulthood, ensuring the long-term reliability and stability of the biometric data.
3. *Unforgeable*: Most vascular patterns exist beneath the skin's surface, making them exceptionally challenging to change or replicate. Furthermore, unlike face recognition, vascular patterns cannot be surreptitiously captured.
4. *Contactless capture*: Vascular images are typically

captured using contactless sensors, making them more hygienic and convenient.

5. *Robustness*: Studies have shown that vascular modalities such as finger-vein recognition remain highly accurate even under various types of distortion, such as motion blur, defocus, sensor aging, and pixel defects.

Nonetheless, vascular biometric templates need protection for security and privacy reasons, not to mention the *irrevocability* problem of raw biometric references, meaning that if the raw biometric information is compromised, unlike a password, the corresponding biometric traits such as vein patterns cannot be changed. To mitigate such issues, one may use biometric template protection methods to encrypt or transform the raw biometric templates into a protected form. These methods aim to protect biometric data from unauthorized access, misuse, and theft [27].

These biometric template protection methods have been developed to address these concerns. These protection methods can be broadly classified into (i) cryptographic and (ii) non-cryptographic methods. Cryptographic methods, such as homomorphic encryption, fuzzy encryption, secure multi-party computation, and the often-used standard encryption methods employ mathematical algorithms to encrypt biometric templates, ideally while preserving the original distance relationships between genuine and impostor comparisons so that the matching won't need to happen in the decrypted space. Non-cryptographic methods, such as feature transformation techniques, modify the biometric templates to make them less recognizable while trying to retain their discriminative power.

This paper presents a novel approach to secure private key generation for authentication systems. Our approach leverages Support Vector Machines (SVMs) as a replacement for storing biometric templates while achieving high accuracy comparable to that of deep feature matching. Additionally, we explore various levels of deep feature extraction and fusion techniques to provide the best input for the private key generation method and introduce the usage of clustering techniques such as k-means to keep the identity of individuals anonymous. This eliminates the risk of data breaches and enhances user privacy. Our unique contributions are particularly significant in addressing the key challenges in biometric authentication systems, such as privacy, security, and accuracy.

The rest of the paper is structured as follows. In Section 2, we provide an overview of the related work in the field of biometric authentication and key generation. Section 3 describes our proposed methodology in detail, including our feature selection process, matching methods as a baseline model, and the use of SVMs for secure template replacement. We also explain our key generation method involving the clustering of imposters and its impact on system accuracy. Finally, we conclude our paper in Section 4, summarizing our contributions and highlighting the significance of our proposed approach.

2. Related Work

Proper design and implementation of biometric systems require addressing concerns not only related to the (unbiased) accuracy but also the privacy and security of biometric data [22]. The literature suggests various approaches, such as biometric cryptosystems and cancellable biometrics, to improve the security and reliability of biometric systems and address susceptibility to attacks, scalability, and ethical concerns regarding the collection and storage of biometric data, some of which are presented below.

Locality-sensitive Hashing. Locality-sensitive hashing (LSH) is an approximate nearest neighbor search technique in high-dimensional spaces. The idea behind LSH is to hash the data points so that similar points are mapped to the same bucket with a high probability. In contrast, dissimilar points are mapped to different buckets with a high probability. [11]. LSH may be used to derive binary keys from biometric templates [21].

Support Vector Machines. A Support Vector Machine is a supervised machine learning algorithm that finds a hyperplane in a high-dimensional space to separate data points of different classes [15]. The goal of SVM is to find a hyperplane such that it adopts the maximum distance from the closest data points of two classes. SVMs have been used in various capacities in the design of biometric systems, including key generation [20].

K-Means Clustering. K-means is a popular unsupervised clustering algorithm used in machine learning and data mining to group similar data points together. The k-means algorithm aims to iteratively minimize a distance metric, such as the sum of squared distances between each data point and its assigned cluster center. K-means clustering may be used to assist in the binarization of biometric data [5].

Error Correction Methods. Error correction codes such as Reed-Solomon [14] (RS) algorithm can be used to detect and correct errors in binary data. RS is based on a finite field where a message is first converted into a polynomial and then encoded by adding redundancies. During transmission, if some symbols are lost or corrupted, the receiver can use these redundant symbols to detect and correct errors in the message. The number of redundant symbols added to the polynomial depends on the desired level of error correction. Such methods may be used to help with biometrically-enabled cryptosystems [16].

Biometric Cryptosystems. Biometric Cryptosystem (BCS) combines characteristics of both the fields: biometric and cryptosystem, where biometric provides authentication and cryptosystem imparts security [13]. One of the main advantages of a BCS is that it is highly resistant to spoofing or impersonation attacks since the cryptographic key is generated directly from the user's biometric data and cannot be easily replicated or forged. This makes BCS systems ideal for use in high-security applications such as financial transactions, government systems, and military installations.

Fuzzy Vault. The fuzzy vault scheme is a biometric template protection method that secures a secret key and an unordered set of biometric features in an indivisible vault by encrypting and decrypting them. It generates a user-specific key and constructs a polynomial of degree 'L' by applying an error-correcting code to the encoded key as coefficients. Chaff points are added to the set to hide genuine points, and during authentication, a probe biometric template is presented to recover the key. Successful reconstruction of the polynomial and recovery of the key requires the identification of at least 'L+1' genuine points from the vault. [17].

Secure Sketches. Secure sketches encode secret messages in noisy channels by adding redundancy to the original message. They enable the recovery of lost or corrupted parts during transmission. Biometric data such as fingerprints, facial features, and iris scans can be protected using secure sketches with the use of a secret key to transform the data. This allows secure transmission over insecure channels while maintaining privacy and security. [4].

2.1. Biometric Terms

Helper Data. Helper Data refers to the additional data used to assist the process of biometric cryptography or protected biometric matching without revealing any sensitive information. Helper data is usually stored along with protected data and is used during the biometric verification phase. The use of helper data in secure biometric matching has been shown to improve the accuracy of the secured system [7].

ROC Area Under the Curve. The AUC of a ROC (Receiver Operating Characteristic) curve represents the overall ability of a classifier to distinguish between positive and negative classes. An AUC close to 1 is indicative of high accuracy, and an AUC of 0.5 depicts the worst-case scenario (random decisions).

d-prime. ”d prime” (d') is a statistical measure used to show the separability of genuine imposter score distributions. It is calculated by taking the difference between the means of the two distributions and dividing it by the corresponding standard deviations, with a higher d' value indicating better distinguishability between genuine and impostor scores and, therefore, better accuracy. Unlike ROC AUC, d' does not saturate.

Equal Error Rate. EER[1] is a scalar representation of the accuracy of a binary classifier. It represents the threshold where the false rejection and false acceptance rates are equal.

Rank-k Accuracy. The rank-k accuracy is an important evaluation metric for biometric identification systems (as opposed to the earlier mentioned verification metrics). It measures the percentage of times the correct match of a probe against a gallery is within the top k results of the system's ranked output[2].

3. Methodology

3.1. Dataset

For this study, we used two publicly available finger vein datasets[3]: the SCUT FV Presentation Attack Database (SCUT FVD) [18] and the UTFVP Finger vein Database [25, 26]. The SCUT FVD dataset contains vein images from 100 participants, with six fingers and six vein images per finger, for a total of 600 unique fingers and 3600 captures. On the other hand, the UTFVP Finger vein database contains 1440 images from 60 subjects, with 6 fingers and 4 vein images per finger, for a total of 360 unique fingers. By using these datasets in our experiments, we aimed to evaluate the effectiveness of our proposed method in a realistic and challenging cross-dataset scenario.

3.2. Data Preprocessing

In this study, we followed standard pre-processing practices to prepare our image data for use with deep learning models. The data pre-processing steps include color jittering and image resizing.

1. The color jittering step applies random adjustments to the brightness, contrast, saturation, and hue of the input image, which can increase the diversity of the training data and improve the model's ability to generalize to new data.
2. The resizing step ensures that the input images are all of a consistent size matching our deep learning models' input volume, which is 224x224.

[1] EER = x; where FAR(x) == FRR(x)

[2] For instance, an identification system with a rank-5 accuracy of 95% means that the correct match is among the top 5 results 95% of the time

[3] We obtained permission to use both datasets

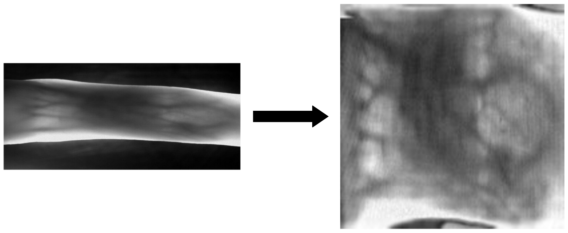


Figure 1. A sample finger vein image from the UTFVP dataset before and after processing using the pipeline in Fig. 2.

The SCUT FVD dataset contains images already processed to obtain the region of interest (ROI), which is not the case for images from UTFVP. Fig.1 shows the finger vein image from the UTFVP dataset before and after processing. In their preprocessed state, the bright band of light is the ROI, and the outer dark space is the experimental area. In the image processing pipeline, the information is extracted from the ROI of the finger without losing any information.

The finger edges in raw images were detected using the Canny algorithm [23]. Linear regression is used to obtain the ROI by fitting it over the boundary edges. An affine transformation is then applied to resize to 224 x 224 images. Gaussian blur is applied to remove false noise, and histogram equalization to enhance vein pattern visibility.

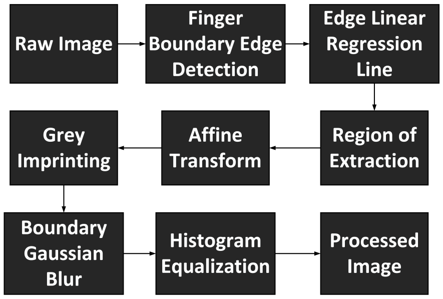


Figure 2. Image processing pipeline.

3.3. Classification Methods

We used four different pre-trained CNN models for deep feature extraction: ResNet50, VGG19, GoogLeNet, and EfficientNet, pre-trained on ImageNet. This method involves replacing the classification head of a pre-trained model, followed by fine-tuning the transferred model on the SCUT FVD dataset. The dataset is divided such that 4 images from each subject are utilized for training, while the remaining 2 images are reserved for subject-dependent testing. The feature extraction step involves feeding input images through the fine-tuned CNN model and extracting the output from the intermediate layers. In addition to the 'flatten' layer from all four models, we also extracted features from several deeper layers to evaluate the information contained in those layers. Specifically, for ResNet50, we extracted features from 'layer2.3.add', 'layer3.5.add', and 'layer4.2.add'. For VGG19, we extracted features from 'classifier.4', 'classifier.5', and 'classifier.6'. For GoogLeNet, we extracted features from 'inception4e.cat', 'inception5a.cat', and 'inception5b.cat'. Finally, for EfficientNet, we extracted features from 'features.5.2.add', 'features.6.3.add', and 'features.8.0'.

ArcFace Loss Function. ArcFace is a loss function that aims to improve the softmax loss function, especially for biometric applications such as face recognition using an angular similarity such as cosine to produce match scores. Unlike the softmax loss function, ArcFace considers the angular relationship between different classes and assigns more importance to the correct classification of similar classes [3]. The ArcFace loss function maps the features extracted from the images into a hypersphere and then applies a margin-based penalty to the features corresponding to the incorrect class. This results in a better-defined decision boundary and more accurate classification results in biometric recognition tasks.

Feature Extraction. The 'flatten' layer is a commonly used layer in deep learning models for feature extraction [8]. It usually carries an information-dense representation of class-specific features from its preceding convolutional layers. The reason for also extracting features from deeper layers of the CNN models was to evaluate the information contained in those layers and to determine whether that information can also improve the model's performance for the task at hand. Deeper layers of CNN models typically capture a different abstraction of the higher-level features that may improve the performance of the system [2].

Fusion. Feature fusion [28] and score-level fusion [1] are two commonly used techniques in machine learning in general and biometrics, in particular, [24] for improving accuracy and robustness. Feature fusion involves combining the features extracted from multiple sources, layers in our case, to improve the accuracy of a classification model. This is usually achieved by concatenating the features to combine the features, possibly after normalization and shortening. On the other hand, score fusion involves combining the multiple match scores by non-learning methods such as sum rule (averaging), product rule, minimum rule (akin to fuzzy AND), maximum rule (akin to fuzzy OR); or by stacking a meta-classifier on top to learn a bespoke fusion function. Fusion leverages the complementary information and strengths of different representations and models to improve the overall performance and robustness of the biometric system.

Dimensionality Reduction. Due to the high dimensionality of the features extracted from deeper layers, the computational cost of performing research operations can be substantial. Specifically, following feature concatenation, the dimensionality of the resulting features could be much higher and on the order of 10^6 for larger CNN layers. Consequently, there is a need for dimensionality reduction to efficiently handle such large and complex volumes. A common method we implemented in this study is principal component analysis (PCA). PCA transforms the high-dimensional data into a lower-dimensional subspace by linearly projecting the data onto orthonormal basis vectors (i.e., principal components) that capture the most significant variations in the dataset [29]. The eigenvalue normalization further helps the feature-level fusion. The dimensionality reduction is achieved by setting a goal for the preserved variance, e.g., 95% of the total feature set variance being explained by the resulting components.

Matching Techniques. Cosine similarity-based matching is a technique commonly used in many biometric methods [10] that uses deep features to compare the similarity between enrollment and verification templates.

Multi-Matching. Multi-matching is a multi-sample fusion method used by many successful biometric implementations. It allows a user to be matched against more than one enrollment (or verification) template. For instance, the user may produce multiple templates at the time of enrollment to allow for natural variabilities in their biometric samples to be captured and stored. During verification, the presented biometric is matched against the plurality of the claimed stored templates (the template bank). The resulting multiple-match scores are then fused to yield the final result, which is usually superior in terms of accuracy and robustness, albeit at the expense of added computational and storage overheads.

Comparison between ArcFace and Cross-Entropy. For the pre-trained ResNet50, we found the match results (both for single and multi-sample) to be very close. For training and subject-dependent testing, ArcFace was in the lead. Subject-independent testing showed slightly better results with the cross-entropy loss. The subject-independent testing protocol used 100 identities from SCUT FVD that were never used for any training or validation. Among the remaining 500 identities, 4 images per identity were used for model training, and the remaining 2 were used for subject-dependent testing. In the face of the above results, we moved forward with ArcFace. It is conceivable that with larger datasets, we could see better differentiation between these two loss functions.

Feature Extraction Layer[4]	SCUT FVD	UTFVP
layer2.3.add	9.447	2.382
layer3.5.add	7.0085	2.662
layer4.2.add	0.5064	0.506
flatten	0.09408	0.024
Fusion of Features	0.6474	0.496

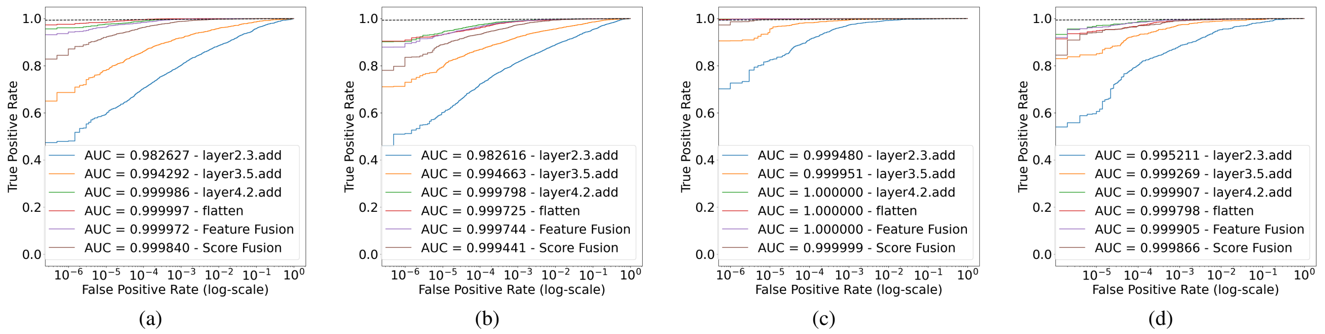
Table 1. All-impostor subject independent test accuracy(in %) for various features extracted using fine-tuned ResNet50 with ArcFace, measured at the training set EER threshold.

Subject-Independent Results. Our subject-independent testing uses a model that was fine-tuned on a subset of data with 500 identities but tested on an exclusive subset of 100 identities that had not been exposed during the model's transfer learning and fine-tuning. Subsequently, we calculated the cosine similarities for the features of the exclusive 100-identity subject-independent test set. Any cosine similarity values exceeding the established equal error rate (EER) threshold acquired from the *training* set were considered as positively authenticated. For the identification task, this means that the rank-k metrics were calculated after passing such a matching threshold. For an all-impostor subject-independent matching, we anticipate the rank-k accuracies to be very low and thus can use this test to glean the quality of different features, as depicted in Table 1.

3.4. Template Protection Method

In this section, we describe our proposed template protection method that essentially uses a binarizing classifier on top of the best features discovered during the above-mentioned section and learns to produce a binary code. The resulting binarization turns out to be more stable than many comparable methods since distance-preserving hashes may produce inconsistent bits given the natural variations of biometric samples and features. Stabilizing these wavering bits by way of error-correcting codes may reduce the entropy and, thus, the uniqueness of the resulting binary keys, as we shall see later. Instead of saving the biometric template, we propose saving the aforesaid classifier parameters as helper data. This basic approach enables our biometric system to produce unique and revocable private keys from the earlier-discussed deep features. Such keys, after provisioning, can be used in a host of cryptographic operations such as unlocking a vault, private-public key exchanges, distributed ledgers, or deciphering messages. We take this approach further by enhancing the privacy and security of the system using a privacy-preserving derived impostor class for the aforesaid

[4] All features are from the 'flatten' layer after PCA processing with 95% total variance retained



(a) (b) (c) (d)

Figure 3. (a) Training single match, (b) Subject dependent testing single match, (c) Training multi-match, (d) Subject dependent testing multi-match; ROC curves on a logarithmic scale for ResNet50 trained using ArcFace using a subject dependent protocol.

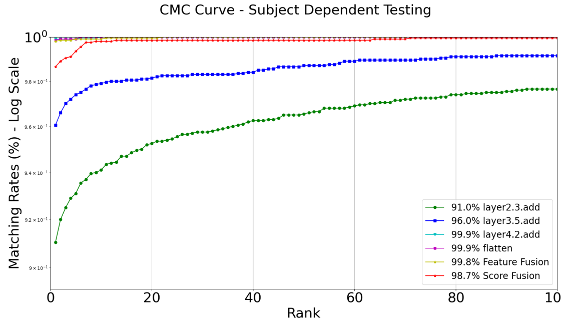


Figure 4. Open-set CMC curve on a logarithmic scale for ArcFace fine-tuned ResNet50. The legend shows Rank-1 accuracies for each layer.

binarizing classifier. Finally, we create a self-referencing version of this method that generates the equivalent of the enrollment key on the fly for no-reference secure and private biometric matching.

Locality Sensitive Hashing as a baseline. We employed LSH to create binarizing random planes in two experiments. In the first experiment, 500 random planes were generated for all training data points (subject-agnostic LSH binarization). We observed the distribution of the Hamming distances of genuine imposter keys as a measure of goodness [6]. We then pruned these planes down to 64 based on the d' values of the bit locations.

In the second experiment, planes were generated for each identity (subject-specific LSH binarization), with a 32-bit random bit sequence assigned during the enrollment and the selected 32 planes (with best separability between the chosen ID vs. rest based on d') forced to generate the corresponding bit sequence by manipulating coefficient signs. For verification, the claimed identity's plane set was used to regenerate the bit sequence using the verification feature vector.

While the results were rather promising for the subject-specific LSH plane banks, there was a high rate of false rejects. To address this issue, we used a Reed Solomon error correction with the capability of correcting up to 2 bits, which we determined from sample Hamming distances. The resulting syndrome bits were saved alongside the LSH planes to be recalled during verification. However, it was observed that some imposter bit sequences were corrected into false matches, leading to an increase in false accept rates. To remove the need for storing the syndrome bits in their original format and to partially tackle the earlier mentioned problem, we created a second LSH bank for each identity at the time of enrollment that was programmed to generate for the syndrome part of the Reed Solomon code. This helped with false accepts, reducing the chances of impostor key correction given that the RS syndrome also has to be recovered by the impostor.

Key generation using SVM banks. This was achieved by having an SVM classifier for each of the 32 bits in the randomly generated private key for each identity at the enrollment phase. During the SVM training process, we utilized features from both genuine and imposter classes, with the selection of the target class being dependent on the specific bit being trained. In order to effectively train the non-target class, we randomly selected a set of (I_N) number imposters and used their corresponding features using algorithm 1.

To illustrate, consider generating an SVM classifier for a specific subject, '*R*,' using a dataset consisting of {0, 1, 2, ..., R, ..., N}. During training, we designated features of R as the target class data and used data points from a random selection of imposters as the non-target class. Depending on the expected bit (i.e., 1 or 0), we set data of 'R' as the target class for the corresponding SVM classifier and the non-target class as the opposite class (i.e., 0 or 1). Various SVM kernels, including linear, quadratic, and cubic, were experimented with. However, as the degree of the kernel polynomial was increased, the classifiers began to overfit on the training data, which led to higher error rates on the test data. As a result, the study proceeded with the linear

Algorithm 1 Algorithm for selecting imposter sequences

```
function IS_CLOSER(list, num, padding)
    for element in list do
        if abs(element - num) $\leq$ padding then
            return True
        end if
    end for
    return False
end function
Shuffle the list of SUBJECT_LABELS
Initialize empty list SEEN
Initialize empty list $SEQ_{LIST}$
Initialize integer $I_{PAD}$ based on number subjects
Set counter N to 0
while $N < 32$ do
    Generate a random integer $I_R$ such that it is at least $I_{PAD}$ indices away from the beginning and end of imp_list
    if not is_closer(SEEN, $I_R$, $I_{PAD}$) then
        Append $I_R$ to SEEN
        Append the sub-sequence() of length 2*$I_{PAD}$ centered at $I_R$ to $SEQ_{LIST}$
        Increment N by 1
    else
        Generate a new random integer for $I_R$
    end if
end while
return $SEQ_{LIST}$
```

kernel.

By storing the parameters of these SVM classifiers, we reconstruct and apply them to the verification features and get back the original 32-bit sequence, obviating the need to store either the template of the secret private keys, thereby ensuring the confidentiality and security of the biometric data. The key generation results of these SVM banks, as judged by the genuine and impostor Hamming distances of the generated keys vs. the targets, were much more accurate when compared to the LSH method. This should not come as a surprise, given that the SVMs learn the proper bit sequence using maximum margin classification compared to the random projections of the LSH method. It should also be noted that due to the accuracy of the SVMs, we did not use RS error correction. We should add that we can extend this methodology to generate longer keys, such as 64 or even 128-bit sequences. However, given the perceived accuracy of this pipeline, we decided that 32 bits can adequately accommodate the entropy of the system.

SHA 256 encoded reference for key-based verification. So far, the proposed system generates private keys from the presented biometric samples for ensuing operations such as deciphering a text or unlocking a key vault, but due to a lack of stored biometric reference, one cannot readily carry out verification matching tasks. One way to do so is to treat the generated private key as a password and compare its hash with the stored hash of the original key for security purposes. We implemented this additional hashing layer of protection by encoding the assigned 32-bit sequence using an SHA 256 hash function [19]. Subsequently, the first 32 bits of the resulting hash value were utilized as the shortened hash for the next steps. We note that inverting the SHA 256 hash function is considered infeasible thus far. During verification, the presented biometric sample for claimant of identity ‘S’ undergoes feature extraction and subsequently classification via SVM bank of classifiers specifically trained for generating the key for identity ’S’. The generated bit sequence undergoes SHA 256 encoding, and the first 32 bits of the produced hash are compared with the stored hash for final verification.

As the next step for the evolution of this mechanism, we decided not to even store the hash of the intended key and use another SVM bank to produce the reference hash out of an abundance of caution. In this iteration, we utilized a second set of SVMs trained to convert the presented biometric feature to the intended 32-bit hash. Obviously, the training of this secondary SVM bank occurs during the enrolment and the provisioning of the original key, where the parameters of this SVM bank are stored alongside the first banks as augmented helper data. The resulting bit sequence from this secondary bank is then compared to the hash of the first bank’s for the verification of the identity ‘S.’

3.5. Clustering data for preserving privacy

The SVM-based private key generation method described in the earlier section requires a token sample of impostor features to train the SVM banks, meaning that such impostor feature sets need to be provided during the provisioning of the private keys for a new user. Such practice may create privacy concerns. Therefore, we created synthesized impostor feature sets for this purpose using a clustering technique. To do so, the imposter samples for a new identity’s SVM training are grouped into 32 clusters. Each cluster has a centroid that represents the average of all the data points in that cluster, masking the original features. These centroids are then matched against the new user’s genuine samples to train their key-generating SVM banks. As an added bonus, this clustering also improved the accuracy of the SVM classifiers, possibly by averaging out the noise.

In this final form of our proposed method, to evaluate the accuracy of key-based verification in a multi-match context, we consider the following cases based on the Hamming distances of the comparisons (two in our case):

Cosine Similarity	d'	AUC (%)	EER (%)
Train	36.68	99.9997	0.053
Test	26.98	99.9725	0.604
MM Test	28.86	99.9798	0.453

SVM	d-prime	AUC (%)	EER (%)
Train	107.26	99.5383	0.462
Test	52.19	99.0684	0.932
MM Test	81.21	99.717	0.491

SHA-SVM	d-prime	AUC (%)	EER (%)
Train	111.67	99.5563	0.444
Test	53.05	99.0832	0.918
MM Test	76.64	99.6332	0.576

32 Clusters	d-prime	AUC (%)	EER (%)
Train	36.48	99.9868	0.0132
Test	36.55	99.9003	0.673
MM Test	40.38	99.974	0.742

64 Clusters	d-prime	AUC (%)	EER (%)
Train	70.54	99.992	0.008
Test	61.57	99.908	0.716
MM Test	73.41	99.9813	0.631

Table 2. Comparing all the experiments for the 'flatten' layer features from ResNet50 fine-tuned with ArcFace. MM stands for multi-match. The key-generating ROCs use Hamming distances in lieu of match scores.

Case - I: If any one of the key comparisons has a zero hamming distance, the verification attempt is considered to be successful. This is akin to OR logic and more prone to acceptance.

Case - II: Only if all the key comparisons have a zero Hamming distance the verification attempt is considered to be successful. This is akin to AND logic and more prone to rejections.

Case - III: If there is a mismatch between the multi-match Hamming distances, i.e., comparison with the first enrollment sample generates a zero Hamming distance, but the comparison with the second enrollment sample generates a non-zero Hamming distance, we consider this case as a no-decision and for instance ask for a recapture, in which case we may call this situation a Failure to Capture, or FTC, as well. These FTCs for 32 cluster impostors were 23.41% , and 13.35% for 64 cluster case.

3.6. Comparison of the Results

The overall results presented in table 2 show that the privacy-preserving impostor clustering method yielded relatively high AUCs for both the training and testing, suggesting that it is effective in helping the key-generating SVMs distinguish between genuine and impostor pairs. Nevertheless, it is worth noting that the d' decreases for the clustering method, something that multi-match seems to be able to compensate for in part.

4. Discussion

Our presented approach first focuses on deriving the best deep features by evaluating different models, feature presentations, loss functions, and fusion techniques. Specifically, we examine the performance of deeper layers of CNN models in comparison to the 'flatten' layer, which is commonly used for feature extraction. Our results demonstrate that vascular features extracted from deeper layers of CNN models can come close to those extracted from the 'flatten' layer, but their higher dimensionalities need to be mitigated.

We used SVM classifiers for key generation to provide several advantages over traditional cosine similarity-based methods. First, our proposed method generates a stable key without loss of accuracy out of the biometric features that may even achieve better EER and d-prime values when compared to traditional matching methods. Such biometrically generated keys can enable a host of security and cryptographical applications. Furthermore, this highly customizable approach allows for revocability by simply replacing the SVM banks. Most importantly, it is the SVM parameters and not the original biometric template that is being stored, mitigating many privacy and security concerns.

The addition of SHA hashing and creating a secondary bank of SVMs to produce the SHA hash of the first SVM bank allows the proposed approach to create verification results by comparing these keys post-hashing and generating a matching decision without the need for a stored biometric template. In order to improve the privacy aspect of the key-generating SVMs during their training, we introduced a clustering method to generate anonymized impostor data without directly sending over real biometric templates or impacting accuracy.

5. Acknowledgements

The research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-21-2-0252. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes, notwithstanding any copyright notation herein. Dr. Derakhshani is also a consultant for Jumio.

References

[1] K. Aizi and M. Ouslim. Score level fusion in multi-biometric identification based on zones of interest. *Journal of*

King Saud University-Computer and Information Sciences, 34(1):1498–1509, 2022.

[2] A. Boyd, A. Czajka, and K. Bowyer. Deep learning-based feature extraction in iris recognition: Use existing models, fine-tune or train from scratch? In *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–9. IEEE, 2019.

[3] J. Deng, J. Guo, N. Xue, and S. Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4690–4699, 2019.

[4] B. Fuller, X. Meng, and L. Reyzin. Computational fuzzy extractors. *Information and Computation*, 275:104602, 2020.

[5] J. García, B. Crawford, R. Soto, C. Castro, and F. Paredes. A k-means binarization framework applied to multidimensional knapsack problem. *Applied Intelligence*, 48:357–380, 2018.

[6] S. Grabowski and T. M. Kowalski. Algorithms for all-pairs hamming distance based similarity. *Software: Practice and Experience*, 51(7):1580–1590, 2021.

[7] M. I. Hashem and K. Alibraheemi. Literature survey: Biometric cryptosystems based on fingerprint processing techniques. In *2022 International Conference on Data Science and Intelligent Computing (ICDSIC)*, pages 198–201. IEEE, 2022.

[8] H. Heidari and A. Chalechale. Biometric authentication using a deep learning approach based on different level fusion of finger knuckle print and fingernail. *Expert Systems with Applications*, 191:116278, 2022.

[9] B. Hou, H. Zhang, and R. Yan. Finger-vein biometric recognition: A review. *IEEE Transactions on Instrumentation and Measurement*, 2022.

[10] N. Ibtehaz, M. E. Chowdhury, A. Khandakar, S. Kiranyaz, M. S. Rahman, A. Tahir, Y. Qiblawey, and T. Rahman. Edith: Ecg biometrics aided by deep learning for reliable individual authentication. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 6(4):928–940, 2021.

[11] O. Jafari, P. Maurya, P. Nagarkar, K. M. Islam, and C. Crushev. A survey on locality sensitive hashing algorithms and their applications. *arXiv preprint arXiv:2102.08942*, 2021.

[12] A. K. Jain and K. Nandakumar. Biometric authentication: System security and user privacy. *Computer*, 45(11):87–92, 2012.

[13] P. Kaur, N. Kumar, and M. Singh. Biometric cryptosystems: a comprehensive survey. *Multimedia Tools and Applications*, pages 1–56, 2022.

[14] S. Kumar, D. Sharma, and A. Payal. Performance enhancement of multi channel multi beam fso communication link with the application of reed solomon codes. *Optical and Quantum Electronics*, 54(11):740, 2022.

[15] W. S. Noble. What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567, 2006.

[16] G. Panchal and D. Samanta. A novel approach to fingerprint biometric-based cryptographic key generation and its applications to storage security. *Computers & Electrical Engineering*, 69:461–478, 2018.

[17] W. Ponce-Hernandez, R. Blanco-Gonzalo, J. Liu-Jimenez, and R. Sanchez-Reillo. Fuzzy vault scheme based on fixed-length templates applied to dynamic signature verification. *IEEE Access*, 8:11152–11164, 2020.

[18] X. Qiu, W. Kang, S. Tian, W. Jia, and Z. Huang. Finger vein presentation attack detection using total variation decomposition. *IEEE Transactions on Information Forensics and Security*, 13(2):465–477, 2018.

[19] D. Rachmawati, J. Tarigan, and A. Ginting. A comparative study of message digest 5 (md5) and sha256 algorithm. In *Journal of Physics: Conference Series*, volume 978, page 012116. IOP Publishing, 2018.

[20] J. A. Ramírez-Ruiz, C. F. Pfeiffer, and J. A. Nolazco-Flores. Cryptographic keys generation using fingercodes. In *Advances in Artificial Intelligence-IBERAMIA-SBIA 2006: 2nd International Joint Conference, 10th Ibero-American Conference on AI, 18th Brazilian AI Symposium, Ribeirão Preto, Brazil, October 23-27, 2006. Proceedings*, pages 178–187. Springer, 2006.

[21] D. Sadhya, Z. Akhtar, and D. Dasgupta. A locality sensitive hashing based approach for generating cancelable fingerprints templates. In *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–9. IEEE, 2019.

[22] A. Sarkar and B. K. Singh. A review on performance, security and various biometric template protection schemes for biometric authentication systems. *Multimedia Tools and Applications*, 79:27721–27776, 2020.

[23] E. A. Sekehravani, E. Babulak, and M. Masoodi. Implementing canny edge detection algorithm for noisy image. *Bulletin of Electrical Engineering and Informatics*, 9(4):1404–1410, 2020.

[24] M. Singh, R. Singh, and A. Ross. A comprehensive overview of biometric fusion. *Information Fusion*, 52:187–205, 2019.

[25] P. Tome, M. Vanoni, and S. Marcel. On the vulnerability of finger vein recognition to spoofing. In *IEEE International Conference of the Biometrics Special Interest Group (BIOSIG)*, Sept. 2014.

[26] B. T. Ton and R. N. J. Veldhuis. A high quality finger vascular pattern dataset collected using a custom designed capturing device. In *Proceedings of the 2013 International Conference on Biometrics (ICB), Madrid, Spain*, pages 1–5, 2013.

[27] Q. N. Tran, B. P. Turnbull, and J. Hu. Biometrics and privacy-preservation: How do they evolve? *IEEE Open Journal of the Computer Society*, 2:179–191, 2021.

[28] N. Zeng, P. Wu, Z. Wang, H. Li, W. Liu, and X. Liu. A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection. *IEEE Transactions on Instrumentation and Measurement*, 71:1–14, 2022.

[29] B. Zhao, X. Dong, Y. Guo, X. Jia, and Y. Huang. Pca dimensionality reduction method for image classification. *Neural Processing Letters*, pages 1–22, 2022.