On the Statistical Independence of Parametric Representations in Biometric Cryptosystems: Evaluation and Improvement

Riccardo Musto, Emanuele Maiorana, Ridvan Salih Kuzu, Gabriel Emile Hine and Patrizio Campisi
Roma Tre University, Rome, Italy

Keywords: Biometric Recognition, Template Protection, Biometric Cryptosystems, Statistical Independence, Finger Vein.

Abstract: Biometric recognition is nowadays employed in several real-world applications to automatically authenticate legitimate users. Nonetheless, using biometric traits as personal identifiers raises many privacy and security issues, not affecting traditional approaches performing automatic people recognition. In order to cope with such concerns, and to guarantee the required level of security to the employed biometric templates, several protection schemes have been designed and proposed. The robustness against possible attacks brought to such approaches has been typically investigated under the assumption that the employed biometric representations comprise mutually independent coefficients. Unfortunately, the parametric representations adopted in most biometric recognition systems commonly consist of strongly correlated features, which may be therefore unsuitable to be used in biometric cryptosystems since they would lower the achievable security. In this paper we propose a framework for evaluating the statistical independence of features employed in biometric recognition systems. Furthermore, we investigate the feasibility of improving the mutual independence of representations defined through deep learning approaches by resorting to architectures involving autoencoders, and evaluate the characteristics of the novel templates through the introduced metrics. Tests performed using templates derived from finger-vein patterns are performed to evaluate the introduced framework for statistical independence and the proposed template generation strategies.

1 INTRODUCTION

Biometric recognition systems rely on personal characteristics to define unique identifiers, through which a user can be automatically recognized and granted physical or logical access to specific goods or services (Jain et al., 2011). Thanks to the improved comfort and security this technology offers, with respect to traditional recognition techniques using passwords or tokens, biometric applications are frequently encountered in our daily lives, from using fingerprint or face to unlock a mobile device, to employing iris patterns to enter restricted areas in airports.

Nevertheless, it has to be remarked that, along with several advantages, the use of biometric traits for recognition purposes also brings many potential security and privacy issues. Actually, in case a biometric trait is compromised, as it may happen if covertly acquired or stolen by an attacker, it cannot be revoked or reissued, being it an intrinsic and permanent characteristic of its owner. If the compromised trait is employed in different applications, all of them are not secure anymore. Since the number of biometric traits a subject can exploit is limited, losing the possibility of relying on one of them is a serious limitation. As far as privacy is concerned, biometric traits could reveal sensitive information about their owners, which could be exploited for purposes not related to recognition and potentially discriminatory. Exploiting the uniqueness of biometric traits, it could be also possible to track the activities of a subject across multiple applications. Actually, the EU General Data Protection Regulation (GDPR) states that biometric traits are sensitive and personal data, and should be therefore processed ensuring adequate levels of security.

For all the aforementioned reasons, it is therefore extremely important to design and implement biometric template protection (BTP) schemes, with the aim of preserving the identifiers stored in every databases, guaranteeing the desired properties of renewability, security, and performance (Nandakumar and Jain, 2015). BTP approaches have been typically categorized into two major classes: cancelable biometrics (Patel et al., 2015) and biometric cryptosystems (Rathgeb and Uhl, 2011). The former methods typically employ non-invertible functions, whose defining parameters may be made publicly available or not, to

Musto, R., Maiorana, E., Kuzu, R., Hine, G. and Campisi, P.
On the Statistical Independence of Parametric Representations in Biometric Cryptosystems: Evaluation and Improvement.
DOI: 10.5220/0010988300003122
In *Proceedings of the 11th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2022)*, pages 480-487
ISBN: 978-989-758-549-4; ISSN: 2184-4313

transform the original representations into protected ones. The latter ones either directly extract binary keys from the considered biometric traits, or adopt key-binding approaches, exploiting hashing protocols to bind biometric templates with binary cryptographic keys. Biometric cryptosystems typically rely on some public information, known as helper data, during the recognition process.

It is worth mentioning that BTP methods relying on cancelable biometrics are commonly analyzed, in terms of robustness against attacks trying to recover the original information from the transformed one, in quite superficial modalities, due to the difficulty in quantitatively evaluating the achievable non-invertibility, and to the heterogeneity of the proposed methods, which makes it arduous to define general metrics upon which quantifying the provided security. Conversely, rigorous evaluations of biometric cryptosystems security have been often provided in literature (Simoens et al., 2009), with in-depth information theoretic studies specifically dedicated to key-binding approaches, trying to evaluate the amount of information leaked by the stored helper data about the original secret sources (Ignatenko and Willems, 2015).

It has yet to be observed that all the theoretical evaluations so far carried out to evaluate the security of biometric cryptosystems have assumed ideal characteristics for the biometric templates to be protected. Specifically, it is typically assumed that the biometric representations employed in key-binding schemes consist of mutually-independent features, an hypothesis allowing to derive conclusions regarding a single coefficient and automatically extending them to the whole set of available features. Unfortunately, real-world data are not characterized by such an ideal property, and the loss in security of a biometric cryptosystem is greater the further the available data are from the ideal condition. It is therefore required to perform proper investigations to evaluate the extent of the differences between ideal and real qualities of the employed biometric representations, having in mind their exploitation in biometric cryptosystems. Furthermore, it would be desirable to design strategies for generating biometric representations having properties as close as possible to the ideal ones, while preserving their discriminative capabilities.

Within the context of biometric cryptosystems, the present paper addresses the aforementioned aspects, performing an analysis regarding the statistical independence of the features comprised within biometric templates. Specifically, a novel framework, through which it is possible to derive quantitative evaluations on the independence of the considered representations, is proposed in Section 2, and applied to analyze templates obtained applying deep learning strategies to biometric data. In more detail, finger-vein patters are considered as biometric traits in this study. Furthermore, approaches relying on autoencoders are employed with the aim of improving the non-ideal characteristics of the considered templates, as described in Section 3. The usability of the generated templates within a biometric cryptosystem is evaluated considering a recently-introduced key-binding scheme with zero-leakage capabilities (Hine et al., 2017), summarized in Section 4. The statistical independence of the considered representations is tested through the proposed metrics in Section 5, where their effectiveness in terms of security and recognition rates is also evaluated. Conclusions are eventually drawn in Section 6.

2 FRAMEWORK FOR THE ANALYSIS OF STATISTICAL INDEPENDENCE

The proposed framework for the evaluation of statistical independence in biometric templates relies on the Hilbert-Schmidt Independence Criterion (HSIC) statistical test (Gretton et al., 2007), described in Section 2.1. The information gained running HSIC tests are then processed by methods derived from graph theory (Bondy and Murty, 2008), as detailed in Section 2.2, to provide quantitative metrics about the independence of the considered coefficients. In the following discussion, it is assumed that a dataset $\mathcal{D}$ of biometric templates, collected from u subjects for a total of n samples, each expressed as a feature vector with length m, is available for the conducted statistical analysis. The considered data are arranged as an $n \times m$ matrix, with each row being a biometric template.

2.1 HSIC Statistical Test

Given two random variables $\mathcal{X}$ and $\mathcal{Y}$, the HSIC test estimates the squared Hilbert-Schmidt norm of the population of interest, that is, HSIC$(\mathbf{P}_{xy}, \mathcal{F}, \mathcal{G})$, where $\mathbf{P}_{xy}$ is the joint distribution of $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$, and $\mathcal{F}$ and $\mathcal{G}$ are two reproducing kernel Hilbert spaces (RKHS). The null and research hypotheses of the HSIC test are defined as:

$$\mathcal{I}(\mathcal{Z}) : (\mathcal{X} \times \mathcal{Y})^n \mapsto 0, 1, \qquad (1)$$

$$\begin{aligned} \mathrm{H}_0 &: \mathbf{P}_{xy} = \mathbf{P}_x \mathbf{P}_y \\ \mathrm{H}_1 &: \mathbf{P}_{xy} \neq \mathbf{P}_x \mathbf{P}_y, \end{aligned} \qquad (2)$$

that is, the null hypothesis correspond to independent $\mathcal{X}$ and $\mathcal{Y}$ variables.

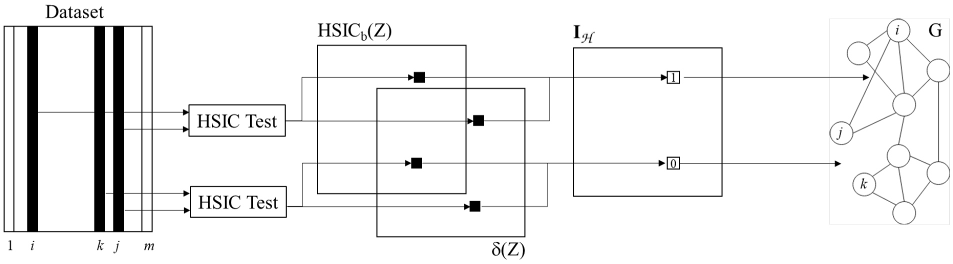


Figure 1: Visual depiction of the proposed framework for statistical independence evaluation.

The desired statistics can be estimated from an observed sample $Z = (X,Y)$ in a biased version as:

$$\mathrm{HSIC}_b(Z) = \frac{1}{n^2}\mathrm{trace}(\mathbf{KHLH}), \tag{3}$$

that is, computing the sum of elements on the main diagonal of a square matrix obtained as the product of the $n \times n$ matrices $\mathbf{H} = \mathbf{I} - \frac{1}{n}11^\top$, $\mathbf{K}$, and $\mathbf{L}$, with the latter two having as elements:

$$\begin{aligned}\mathbf{K}(i,j) &:= exp(-\sigma_x^{-2}||x_i - x_j||^2)\\ \mathbf{L}(i,j) &:= exp(-\sigma_y^{-2}||y_i - y_j||^2), (i,j) = 1,\ldots,n.\end{aligned} \tag{4}$$

Having set the desired value of the significance level α for the upper bound of the type I error, the asymptotic distribution of the empirical estimate $\mathrm{HSIC}_b(Z)$ is derived under H_0, and the quantile $1-\alpha$ of this distribution, indicated $\delta(Z)$, can be used as a threshold to determine the test outcome. Specifically, in case $\mathrm{HSIC}_b(Z) < \delta(Z)$, it is not possible to reject the null hypothesis H_0, thus the two random variables are assumed as independent.

In our framework the HSIC test is performed for every possible pairs of features in the considered m-dimensional biometric representation. An $m \times m$ square and symmetrical independence matrix $\mathbf{I}_{\mathcal{H}} \in \mathbb{Z}^{m \times m}$ can be thus obtained as:

$$\mathbf{I}_{\mathcal{H}}[i,j] = \begin{cases} 1 & \text{if } \mathrm{HSIC}_b(F_i,F_j) < \delta(F_i,F_j) \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where F_i and F_j represent any two features in the set of the m available ones, with samples lying in a $\mathbb{R}^n$ space.

2.2 Graph Theory

Concepts stemming from graph theory are employed to derive quantitative metrics expressing the degree of independence of the coefficients in the considered biometric representation. The computed binary independence matrix $\mathbf{I}_{\mathcal{H}}$ can be in fact interpreted as an adjacency matrix $\mathbf{A}_G$, thus defining an independence undirected graph G, with edges connecting nodes associated to independent coefficients. The overall process performed to derive the desired independence graphs is visually depicted in Figure 1.

In the present paper we consider three different metrics which can be computed from any undirected graph $G = (V,E,\psi)$ having nodes N, edges E, and incidence function ψ, namely *normalized edge count*, *normalized maximum clique size*, and *normalized degree centrality*.

2.2.1 Normalized Edge Count

This metric simply computes the number of edges in the considered graph G, normalized with respect to the maximum number of edges in a complete graph with the same number m of nodes as G, being therefore

$$\mathrm{NEC}_m = \frac{1}{m(m-1)}\sum_{i,j}\mathrm{I}_{\mathcal{H}}[i,j]. \tag{6}$$

The obtained value can be interpreted as a percentage of independent coefficients. Yet, despite its simplicity, this metric is weak and not very exhaustive. As a matter of fact, for example, a value of $\mathrm{NEC}_m = 0.9$ does not mean that 90% of the features are mutually statistically independent, but only that that the independence matrix $\mathrm{I}_{\mathcal{H}}$ contains 90% of unitary entries.

2.2.2 Normalized Maximum Clique Size

A clique of G is defined as a complete subgraph G' of G, such that every two distinct nodes in the clique are adjacent. The clique is said to be *maximal* if it is not a subset of another clique, and *maximum* if it has the largest number of nodes. In the context here considered, the size S of the maximum clique of G is employed as a metric for biometric feature independence, once normalized with respect to the largest possible value, thus obtaining the value $\mathrm{NMCS}_m = S/m$.

Since a clique represents a complete subgraph, this metric gives an effective measure of independence: the features in the maximum clique will be actually mutually independent. However, this metric presents some criticalities. First, the number of elements of the maximum clique may not be large, especially as the number of features increases, returning extremely low independence values that are not suitable for comparison. Moreover, in the event that independence is extremely high, the search for this clique may require a long processing time. Finally, for a given graph there may be multiple maximum cliques with the same number of elements, making it difficult to understand which of them is the best one.

2.2.3 Normalized Degree Centrality

Degree centrality (Freeman, 1978) is a metric which can be associated to each node i in a graph G, by computing as $d_G(i)$ the number of edges incident to the node itself, that is, its number of connections, thus quantifying its importance within the graph. The more important a node is, the more easily it will be crossed by the information flow. A normalized value of degree centrality can be computed dividing it by the maximum feasible degree of the graph, thus obtaining

$$\mathrm{NDC}_m(i) = \frac{d_G(i)}{m-1}. \quad (7)$$

In the ideal case of having all the m available features mutually independent, the maximum clique would correspond to the graph G itself, and every node would have a value of normalized degree centrality equal to 1. In the following, when reporting this metric for the experimental tests described in Section 5, the values computed for each node are organized in a descending order to form a curve, thus better representing the deviation from the ideal scenario of the features under examination. In more detail, the more the curve of the nodes centrality deviates from the ideal one with values only at 1, the less the features are independent. Such representation also allows to give an indication regarding the number of more important nodes, that is, the features independent of most other coefficients.

3 BIOMETRIC REPRESENTATIONS

The proposed framework for statistical independence evaluation has been tested on biometric representations obtained applying deep learning strategies to finger-vein patterns. In more detail, the baseline system here exploited is the one described in (Kuzu et al., 2020a), where finger-vein images have been processed through a convolutional neural network (CNN) derived from DenseNet-161 (Huang et al., 2018), with the addition of a custom set of final layers. In order to generate representations suitable for verification tasks in open-set conditions, the loss function employed for training relies on a cross-entropy function with additive angular margin penalty (AAMP) (Deng et al., 2019). An additional configuration has been here taken into account as baseline, exploiting ResNext-101 (Xie et al., 2017) instead of DenseNet-161, while keeping the same custom set of final layers, and the same size for the generated representations, comprising 1024 coefficients.

In order to verify whether it is possible to improve the independence of the features generated by the aforementioned baseline approaches, we have also evaluated the effectiveness of adding autoencoders in cascade to the employed networks. An analogous approach has been already proposed in (Kuzu et al., 2020b), yet the purpose there was improving the achievable recognition performance, while here the intended goal would be generating biometric representations with more independent features. The rationale behind the proposed approach resides in the fact that autoencoders are typically employed to automatically learn efficient encodings of the processed data, with the aim of reducing the dimensionality of the input representation while keeping all its informative content. While pursuing this target, it is also possible to force the learned representations to assume useful properties, such as sparseness for example. Within the context of the present research, the objective would be therefore to design an autoencoder able to increase the inner independence of the treated representations.

As autoencoder, we have considered the same densely-connected convolutional autoencoder (DCCAE) proposed in (Kuzu et al., 2020b), consisting of a total of 55 layers, with an input layer receiving representations with 1024 coefficients and an inner encoding producing 256 features, and trained it with different loss functions. In more detail, the employed DCCAE has been trained with the aim of minimizing a loss defined as

$$L = L_R + \beta \cdot L_S, \quad (8)$$

where L_R represents the reconstruction loss, computed through the cosine dissimilarity

$$L_R = \frac{1}{B}\sum_{i=1}^{B}[1 - cos(\mathbf{f}_i, \hat{\mathbf{f}}_i)], \quad (9)$$

being $\mathbf{f}_i$ the i-th feature representation generated by the baseline CNN, $\hat{\mathbf{f}}_i$ its counterpart reconstructed by the autoencoder, and B the employed batch size.

In order to evaluate whether distinct approaches may have different effectiveness for the sought target of improving independence, three different loss functions have been here considered for the component L_S, namely *Kullback-Leibler divergence* (KLD), *spectral restricted isometry property* (SRIP) (Bansal et al., 2018), and *DeCov* (Cogswell et al., 2016).

The L_S term based on KLD is defined as:

$$L_S^{KLD} = \sum_{h \in \mathcal{L}} \sum_{j=1}^{N^{(h)}} D_{KL} \left(\rho \| \hat{\rho}_j^{(h)} \right), \; \hat{\rho}_j^{(h)} = \frac{1}{B} \sum_{i=1}^{B} \left[a_j^{(h)} \left(\mathbf{f}_i \right) \right], \tag{10}$$

where $a_j^{(h)}$ is the j-th activation output of the h-th hidden layer of the DCCAE when $\mathbf{f}_i$ is fed as input to the DCCAE, with $j = 1, \ldots, N^{(h)}$, being $N^{(h)}$ the number of activation units in the h-th hidden layer, and $\rho \in [0,1]$ is the sparsity parameter. The set $\mathcal{L}$ represents the layers of the DCCAE dedicated to the inner encoder and decoder, with $\mathcal{L} = \{26-29\}$ for the DCCAE described in (Kuzu et al., 2020b).

The SRIP regularization forces the weights of the network to be near-orthogonal, and it is computed on the weights of each convolutional layer of the proposed DCCAE:

$$L_S^{SRIP} = \sum_{h=1}^{55} \sigma(W^{(h)}) \left(W^{(h)\top} W^{(h)} - I \right), \tag{11}$$

where $W^{(h)\top}$ is a matrix with the weights of the h-th layer, I is the identity matrix and σ is the spectral norm, defined as the largest singular value of $W^{(h)}$.

The *DeCov* loss pushes the network to learn non-redundant representations by minimizing the cross-covariance of hidden activations through a regularization operation. Considering the h-th layer of the employed DCCAE generating the inner encodings, and its activations $a_j^{(h)}$, the desired cross-covariance $\mathbf{C}$ is obtained computing, for all the possible pairs of activations j and k,

$$\mathbf{C}[j,k] = \frac{1}{B} \sum_{i=1}^{B} \left(a_j^{(h)} \left(\mathbf{f}_i \right) - \mu_j \right) \left(a_k^{(h)} \left(\mathbf{f}_i \right) - \mu_k \right), \tag{12}$$

being μ_j the sample mean of activation j over the batch, that is,

$$\mu_j = \frac{1}{B} \sum_{i=1}^{B} a_j^{(h)} \left(\mathbf{f}_i \right). \tag{13}$$

The DeCov loss is then computed as:

$$L_S^{DeCov} = \frac{1}{2} (||\mathbf{C}||_F^2 - ||\mathrm{diag}(\mathbf{C})||_2^2), \tag{14}$$

where $|| \cdot ||_F$ is the Frobenius norm.

4 CONSIDERED BIOMETRIC CRYPTOSYSTEM

In order to verify whether representations generated as described in Section 3 could be actually employed in a biometric cryptosystems while guaranteeing proper recognition and security performance, tests have been conducted considering the zero-leakage BTP scheme presented in (Hine et al., 2017), that is, a code-offset approach inspired by the digital modulation paradigm. As a reference, this method is graphically depicted in Figure 2.

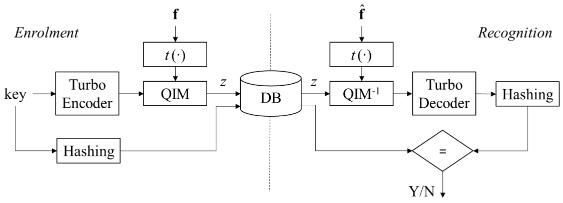


Figure 2: Considered zero-leakage biometric cryptosystem from (Hine et al., 2017).

In brief, it is assumed that a secret binary key and a biometric representation $\mathbf{f} \in \mathbb{R}^m$ are available during enrolment. A point-wise function $t(\cdot)$ is applied to each coefficient of $\mathbf{f}$ in order to derive, from each of them, a variable with a probability density function following a raised cosine distribution, described by a roll-off parameter $\gamma \in [0, 1]$. As shown in (Hine et al., 2017), such transformation guarantees that the stored helper data cannot reveal any information about the employed cryptographic key, thus implementing a zero-leakage protection scheme. The roll-off γ of the employed raised cosine distribution determines the privacy achievable by the proposed scheme, that is, the minimum reconstruction error an attacker can commit when trying to estimate the input biometric sample, given the stored helper data. Specifically, the privacy of the proposed scheme increases with the use of larger values of γ. On the other hand, as shown in (Hine et al., 2017), large γ values reduce the capacity of the considered BTP scheme, that is, the maximum size of the cryptographic key which can be bind with the employed biometric template, thus setting the security of the whole system.

The stored helper data is generated from the input key and biometric template adopting a quantization index modulation (QIM) approach, as $z = [t(\mathbf{f}) - q]_{2\pi}$, where $q \in \{0, 2\pi/P, ..., (P-1)2\pi/P\}^m$, with $P \in \mathbb{N}^+$, is a set of symbols belonging to a phase-shift keying (PSK) constellation of size P, obtained encoding the input binary key with an error correcting code.

The reverse process is performed during recognition, with an inverse QIM applied to the combination of the stored helper data z and the newly acquired biometric representation $\hat{\mathbf{f}}$, after which a soft decoding process with decisions based on log-likelihood ratio (LLR) criteria is performed with the purpose of recovering the originally employed cryptographic key. In case $\mathbf{f}$ and $\hat{\mathbf{f}}$ are similar, the hash of the recovered

Table 1: Comparative analysis of statistical independence of biometric representations obtained using DenseNet-161 as baseline network. Best results reported in **bold**.

Metric	Baseline	ICA	DCCAE		
			KLD	SRIP	DeCov
NEC$_m$	0.152	0.410	0.254	0.250	**0.690**
NMCS$_m$	6	9	7	8	**19**

Table 2: Comparative analysis of statistical independence of biometric representations obtained using ResNext-101 as baseline network. Best results reported in **bold**.

Metric	Baseline	ICA	DCCAE		
			KLD	SRIP	DeCov
NEC$_m$	0.187	0.256	0.221	0.214	**0.415**
NMCS$_m$	8	8	7	7	**10**

key corresponds to that of the secret one, and the subject can be this recognized as a legitimate user.

5 EXPERIMENTAL TESTS

Tests have been performed on finger-vein biometric traits from the SDUMLA database (Yin et al., 2011), comprising data from 106 subjects, with three samples taken from each of three fingers of both hands. The available data have been divided into two disjoint datasets of equal size, one used to train the considered deep learning architectures and the other one to perform the required evaluations. The employed networks have been trained using stochastic gradient descent with momentum (SGDM) and a batch size of 128. The hyperparameters of the DCCAE have been set, for each considered loss L_S, with the aim of guaranteeing the best achievable performance in terms of independence of the generated representations.

The independence metrics proposed in Section 2 have been computed performing HSIC tests with significance level $\alpha = 2.5\%$. A comparative analysis has been conducted in order to evaluate the effectiveness of the methods in Section 3 to generate representations with improved independence, which could be properly used in biometric cryptosystems. In more detail, in addition to comparing the representations generated through the proposed DCCAEs against those obtained using only the considered baseline CNNs, we have also taken into account transformations of the original coefficients obtained through independent component analysis (ICA) (Hyvarinen and Oja, 2000). The considered FastICA approach applies an orthogonal rotation to prewhitened data in order to maximize a measure of non-Gaussianity, used as a proxy for statistical independence.

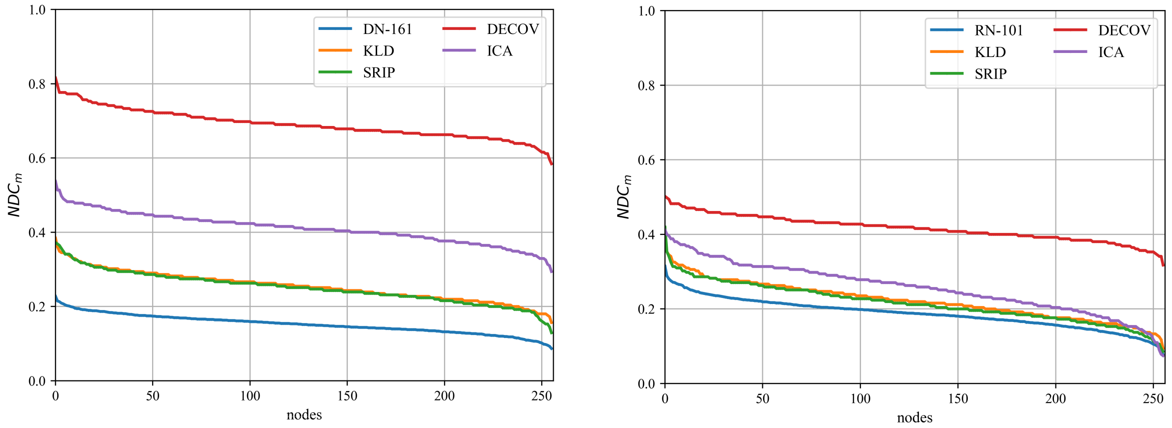
The computed NEC$_m$ and NMCS$_m$ independence metrics are reported in Tables 1 and 2, respectively for biometric representations obtained when using the DenseNet-161 and ResNext-101 baseline networks. The metrics NDC$_m$ are instead reported in Figure 3, which shows the behaviors obtained for all the considered nodes (coefficients), in order to better illustrate the deviation from the ideal conditions with all values set at 1. It is possible to observe that the proposed approaches relying on KLD and SRIP are able to provide an improvement in terms of NEC$_m$ and NDC$_m$ with respect to the use of representations obtained through the baseline networks. The NMCS$_m$ is improved when using DenseNet-161 as baseline. Yet, an ICA transformation guarantee an even further independence improvement. Training a DCCAE with *DeCov* loss represents the best approach to enhance the independence of the considered representation, with notable improvements in terms of NEC$_m$, NMCS$_m$, and NDC$_m$ for templates obtained through both DenseNet-161 and ResNext-101. The gains are especially significant for templates generated with the DenseNet-161 network, with notable increase of independence with respect to ICA too.

In addition to the analysis of the proposed approaches in terms of independence, other aspects relevant to design biometric cryptosystems are also evaluated. The obtained recognition results, in terms of false rejection rate (FRR) and false acceptance rate (FAR), are shown by the detection error tradeoffs (DET) curves of Figure 4, and summarized in terms of equal error rate (EER) in Table 3. It is important to mention that, in this work, the DCCAEs have been trained to maximize feature independence, differently from (Kuzu et al., 2020b) where the networks have been trained to maximize the achievable recognition performance. Therefore, the DCCAEs here employed cannot improve the recognition capabilities offered by the baseline CNN networks as in (Kuzu et al., 2020b). Actually, the use of SRIP and KLD losses guarantee results similar to those of baseline networks, while exploiting the *DeCov* loss worsen the recognition performance achievable with DenseNet-161. As commonly observed in biometric cryptosystem, an improvement in terms of security (here expressed through the achieved independence) can be achieved at the cost of a worsening in terms of recognition performance, with security and recognition involved in a trade-off relationship.

The achievable security is also evaluated by com-

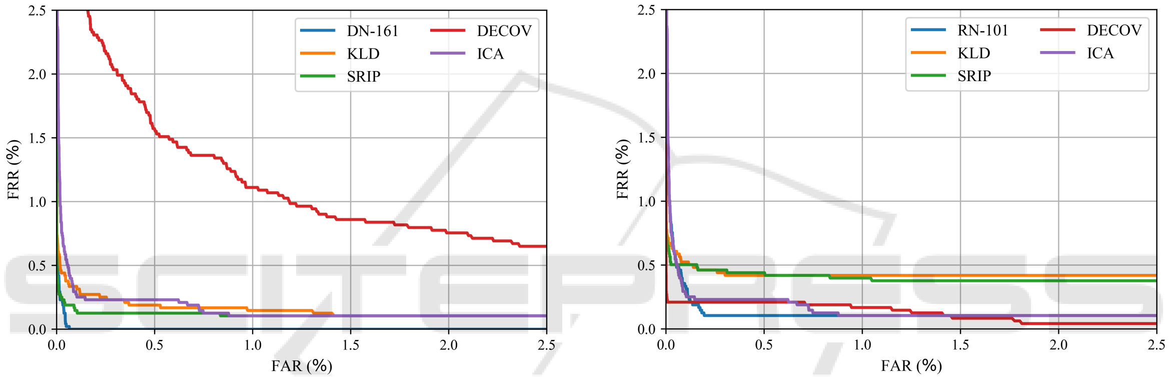
Table 3: EERs (in %) obtained with the considered biometric representations.

Network	Baseline	ICA	DCCAE		
			KLD	SRIP	DeCov
DenseNet-161	0.044	0.586	0.252	0.126	1.072
ResNext-101	0.168	0.230	0.419	0.439	0.209



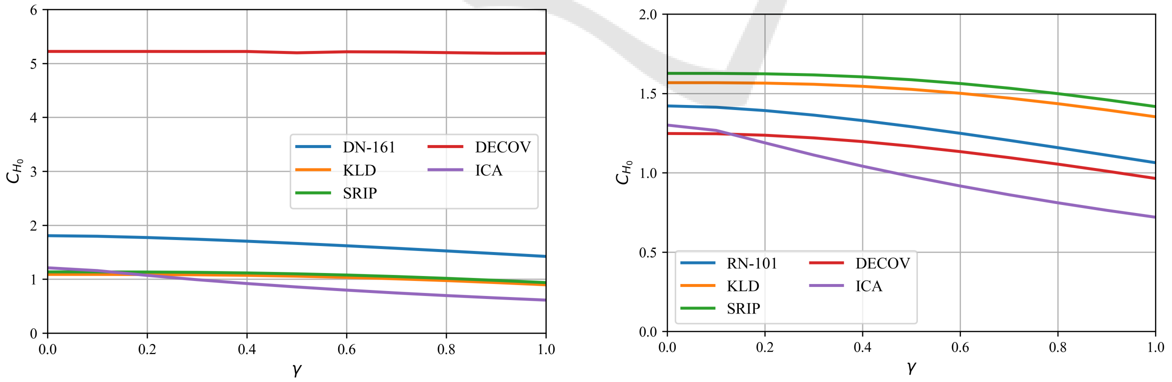
(a) (b)

Figure 3: Normalized degree centrality (NDC$_m$) computed for the considered representations. (a): DenseNet-161 used as baseline; (b): ResNext-101 used as baseline.



(a) (b)

Figure 4: DET curves reporting the recognition performance achievable with the considered representations. (a): DenseNet-161 used as baseline; (b): ResNext-101 used as baseline.



(a) (b)

Figure 5: Average embedding capacity achievable with the considered representations. (a): DenseNet-161 used as baseline; (b): ResNext-101 used as baseline.

puting the capacities of the employed representations, reported in Figure 5 in terms of the average number of bits of a cryptographic key that can be embedded within the coefficients of the employed representation. As detailed in (Hine et al., 2017) and mentioned in Section 4, the achievable capacity depends on the roll-off parameter γ of the raised cosine distribution employed in the considered biometric cryptosystem.

The *DeCov* loss significantly improves the embedding capacity with respect to the baseline system relying on DenseNet-161. It is also worth noting that using an ICA transformation worsens the capacity achievable with representations derived with both DenseNet-161 and ResNext-101. The KLD and SRIP loss functions lead to similar results in terms of channel capacity, with limited improvements with respect to a baseline system only for ResNext-101.

6 CONCLUSIONS

In this paper we have proposed a framework to quantitatively evaluate the statistical independence of features employed within biometric cryptosystems, and used it to analyze the effectiveness of using cascade networks including DCCAEs to produce discriminative and independent biometric representations. Since each of the proposed metric has its own criticalities, it is recommended to evaluate a given biometric representation considering all of them, rather than drawing conclusions based on only one. Further developments could be studied in order to design other metrics with more informative content. For instance, it would be possible to define a new centrality measure, possibly taking into account the difference between the statistics and the thresholds computed by an HSIC test, and associating greater relevance to pair of features with greater differences. This way, evaluations could be performed taking into account the significance of each test, and not only its binary output. Such metric could be also employed to define effective feature selection strategies based on statistical independence. Furthermore, it would be highly desirable to design novel approaches to automatically learn how to generate biometric representations with both discriminative and independence characteristics. To this aim, the proposed metrics could be integrated within the loss functions employed during the learning process.

REFERENCES

Bansal, N., Chen, X., and Wang, Z. (2018). Can we gain more from orthogonality regularizations in training deep cnns? In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 4266–4276, Red Hook, NY, USA. Curran Associates Inc.

Bondy, J. and Murty, U. (2008). *Graph Theory*. Springer Publishing Company, Incorporated, 1st edition.

Cogswell, M., Ahmed, F., Girshick, R., Zitnick, L., and Batra, D. (2016). Reducing overfitting in deep networks by decorrelating representations.

Deng, J., Guo, J., Xue, N., and Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239.

Gretton, A., Fukumizu, K., Teo, C. H., Song, L., Schölkopf, B., and Smola, A. (2007). A kernel statistical test of independence. In *Proceedings of the 2007 Conference on Advances in Neural Information Processing Systems*.

Hine, G., Maiorana, E., and Campisi, P. (2017). A zero-leakage fuzzy embedder from the theoretical formulation to real data. *IEEE Transactions on Information Forensics and Security*, 12(7):1724–1734.

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2018). Densely connected convolutional networks.

Hyvarinen, A. and Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-5):411–430.

Ignatenko, T. and Willems, F. (2015). Fundamental limits for privacy-preserving biometric identification systems that support authentication. *IEEE Trans. on Information Theory*, 61(10):5583–5594.

Jain, A., Ross, A., and Nandakumar, K. (2011). *Introductions to biometrics*. SPRINGER.

Kuzu, R., Maiorana, E., and Campisi, P. (2020a). Loss functions for cnn-based biometric vein recognition. In *European Signal Processing Conference (EUSIPCO)*.

Kuzu, R. S., Maiorana, E., and Campisi, P. (2020b). Vein-based biometric verification using densely-connected convolutional autoencoder. *IEEE Signal Processing Letters*, 27:1869–1873.

Nandakumar, K. and Jain, A. K. (2015). Biometric template protection: Bridging the performance gap between theory and practice. *IEEE Signal Processing Magazine*, 32(5):88–100.

Patel, V. M., Ratha, N. K., and Chellappa, R. (2015). Cancelable biometrics: A review. *IEEE Signal Processing Magazine: Special Issue on Biometric Security and Privacy*, 32(5):54–65.

Rathgeb, C. and Uhl, A. (2011). A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security*, 3:1–25.

Simoens, K., Tuyls, P., and Preneel, B. (2009). Privacy weaknesses in biometric sketches. In *2009 30th IEEE Symposium on Security and Privacy*, pages 188–203.

Xie, S., Girshick, R., Dollár, P., Tu, Z., and He, K. (2017). Aggregated residual transformations for deep neural networks.

Yin, Y., Liu, L., and Sun, X. (2011). Sdumla-hmt: A multimodal biometric database. In Sun, Z., Lai, J., Chen, X., and Tan, T., editors, *Biometric Recognition*, pages 260–268, Berlin, Heidelberg. Springer Berlin Heidelberg.