IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 20, 2021, accepted October 5, 2021, date of publication October 8, 2021, date of current version October 18, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3118830

Feature Fusion Methods for Indexing and Retrieval of Biometric Data: Application to Face Recognition With Privacy Protection

PAWEL DROZDOWSKI, FABIAN STOCKHARDT, CHRISTIAN RATHGEB, DAILE OSORIO-ROIG, AND CHRISTOPH BUSCH, (Member, IEEE)
da/sec–Biometrics and Internet-Security Research Group, Hochschule Darmstadt, 64295 Darmstadt, Germany

Corresponding author: Christian Rathgeb (christian.rathgeb@h-da.de)

This work was supported in part by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE and the European Union's Horizon 2020 Research and Innovation Program through Marie Skłodowska-Curie Grant 860813-TReSPAsS-ETN.

ABSTRACT Computationally efficient, accurate, and privacy-preserving data storage and retrieval are among the key challenges faced by practical deployments of biometric identification systems worldwide. In this work, a method of protected indexing of biometric data is presented. By utilising feature-level fusion of intelligently paired templates, a multi-stage search structure is created. During retrieval, the list of potential candidate identities is successively pre-filtered, thereby reducing the number of template comparisons necessary for a biometric identification transaction. Protection of the biometric probe templates, as well as the stored reference templates and the created index is carried out using homomorphic encryption. The proposed method is extensively evaluated in closed-set and open-set identification scenarios on publicly available databases using two state-of-the-art open-source face recognition systems. With respect to a typical baseline algorithm utilising an exhaustive search-based retrieval algorithm, the proposed method enables a reduction of the computational workload associated with a biometric identification transaction by 90%, while simultaneously suffering no degradation of the biometric performance. Furthermore, by facilitating a seamless integration of template protection with open-source homomorphic encryption libraries, the proposed method guarantees unlinkability, irreversibility, and renewability of the protected biometric data.

INDEX TERMS Biometric identification, biometric template protection, computational workload reduction, indexing, information fusion, face recognition.

I. INTRODUCTION

Personal, commercial, and governmental identity management systems increasingly rely on biometric technologies, which enable reliable recognition of individuals based on highly distinctive characteristics of human beings, *e.g.* face or fingerprints. Applications ranging from personal device access [1], border control [2]–[4], forensic investigations and law enforcement [5]–[7], national ID systems [8], [9], and voter registration [10], [11] benefit from the use of biometrics. The largest systems of this kind enrol hundreds of millions or even beyond a billion enrolled subjects (see *e.g.* [12]), with the global market value of biometric technologies currently estimated to be tens of billions of dollars [13].

As the prevalence, size, and scope of the operational biometric systems increase, the development of technologies which are capable of accurately and efficiently processing biometric data becomes critically important. In the challenging identification and duplicate enrolment check scenarios, where typically an exhaustive search (*i.e.* one-to-many comparison) is needed, solutions which facilitate practical system response times are indispensable. Rather than merely scaling the hardware architecture, which is associated with high monetary costs, algorithmic methods (such as indexing) referred to as biometric workload reduction [14] can be used to speed-up the search queries (and hence reduce the monetary costs). In recent years, strong interest from governmental

The associate editor coordinating the review of this manuscript and approving it for publication was Zhe Jin.

side in such methods has been manifested through numerous benchmarks and competitions [15]–[17].

In addition to the aforementioned practical requirements pertaining to biometric performance and computational efficiency, preventing misuse (*e.g.* privacy violations) of the stored biometric reference data is essential. Existing privacy regulations, *e.g.* the General Data Protection Regulation (GDPR) [18], classify biometric data under ''special categories of personal data'', thus entailing significant responsibilities for the data controllers.

Traditional encryption methods are unsuitable for protecting biometric data, since biometric characteristics exhibit a natural intra-class variance. If traditional cryptographic techniques are applied to biometric templates, said biometric variance prevents a biometric comparison in the encrypted domain. That is, the use of conventional cryptographic methods would require a decryption of protected biometric data prior to the comparison. In contrast, *biometric template protection* [19]–[21] enables a comparison of biometric data in the encrypted domain and hence a permanent protection of biometric data. Biometric template protection schemes use auxiliary data to obtain pseudonymous identifiers from unprotected biometric data. Biometric comparisons are then performed via pseudonymous identifiers while unprotected biometric reference data is discarded [22].

Biometric template protection schemes have hardly been employed in biometric identification systems [14]. One reason for this is that many types of biometric template protection schemes require complex comparison methods which renders them unsuitable for biometric identification (where the workload is dominated by comparison costs). So far, only a handful approaches have combined computational workload reduction strategies with biometric template protection. In the context of face biometrics, those studies have mainly employed cancelable biometrics, *e.g.* [23]–[25]. However, most of those systems still report a degradation w.r.t. biometric performance when benchmarked against unprotected systems. Practically feasible applications of homomorphic encryption in biometric identification systems have likewise been presented [26]–[28]; however, while suffering little to none biometric performance degradation, these schemes also have relied on exhaustive search in a biometric identification scenario and have not considered integration of computational workload reduction such as biometric indexing.

A. CONTRIBUTION AND ORGANISATION

The main contributions of this article are as follows:

- A comprehensive overview and literature survey of works pertaining to (and especially those combining) the areas of information fusion, computational efficiency, and template protection in biometric identification systems.
- A proposal of a multi-stage protected indexing and retrieval system for facial biometric identification based on optimised information fusion and incorporating data privacy-preservation with homomorphic encryption.
- A thorough theoretical analysis and empirical evaluation of the proposed system on a large dataset with state-of-the-art facial recognition systems. Using ISO/IEC IS 19795-1 [29] compliant experimental protocol and metrics, the proposed system is shown to reduce the computational workload of a biometric identification retrieval by approximately 90%, while simultaneously maintaining the baseline biometric performance. Additionally, the possibility of seamless integration of post-quantum-secure homomorphic encryption means that the data security and privacy objectives specified in ISO/IEC IS 24745 [22] are ensured.

The remainder of this article is organised as follows: section II provides relevant background information and an overview of related works. The proposed system is described in section III. The experimental setup and the obtained results are presented in sections IV and V, respectively. Section VI contains concluding remarks and a summary.

II. BACKGROUND AND RELATED WORK

The system proposed in this article combines three research areas within biometrics, *i.e.* information fusion (subsection II-A), computational workload reduction (subsection II-B), and template protection (subsection II-C). This section provides a brief overview of the relevant background information and key related works in those areas.

A. INFORMATION FUSION

Information fusion can be used in order to improve the discriminative power of a biometric recognition system. Referred to as ''multi-biometric systems'', they take advantage of multiple information sources which are combined (fused) in some way. Following fusion categories can be generally distinguished in the context of biometrics [30], [31]:

Multi-type where multiple biometric characteristics (such as facial images and fingerprint scans) are used.

Multi-sensorial where the biometric data acquisition is conducted with diverse sensors providing complementary information (for example, near-infrared and visible-wavelength cameras).

Multi-algorithm where the biometric data is processed utilising several complementary algorithms (for instance, image descriptors based on texture and keypoint information).

Multi-instance where more than one instances of the same underlying type of biometric characteristic are used (*e.g.* the images of right and left iris).

Multi-sample where several biometric samples stemming from one type of biometric characteristic are used (*e.g.* multiple acquisitions of a fingerprint scan with the purpose of detecting reliable regions or assuring the quality and consistency of the acquired data).

Information fusion can occur at different steps of the biometric processing pipeline [30], [31], including:

Sensor where raw data (*e.g.* images acquired by different sensors or multiple samples) is combined before other processing steps [32], [33].

Feature where the extracted feature sets (*e.g.* from multiple samples) are consolidated [34], [35].

Score where the comparison scores computed through different information channels are combined (*e.g.* averaged) [36], [37].

Rank where the orders (ranks) of potential matches between a probe and the enrolment database obtained through different information channels are consolidated [38], [39].

Decision where the decisions (*i.e.* acceptance or rejection) obtained through multiple information channels are combined (*e.g.* by a majority vote) [40], [41].

In the context of this work, fusion of multiple samples (from different data subjects) on feature level are of most interest, as the system proposed in section III is designed to operate at those level of the biometric processing pipeline.

The topic of information fusion in biometrics has been addressed extensively in the scientific literature. In [30], a general introduction to this topic is given, while [31], [42], [43] provide recent and comprehensive surveys of this research area.

B. COMPUTATIONAL WORKLOAD REDUCTION

Maintaining fast biometric identification system response times often requires optimisation or additional investments as the size of the enrolment database increases. The computational costs of the typical, exhaustive search-based, retrieval method tend to grow linearly with the number of enrolled data subjects [44]. Naturally, the expansion of the underlying hardware (*e.g.* by using many servers which facilitate distributing the computations) can be used to maintain quick system response times; however, this solution carries with it high monetary costs, such as the purchase of the equipment, its installation and maintenance, *etc*. While hardware investments are often inevitable, an often overlooked possibility is the optimisation of the underlying software and/or algorithms.

In this context, the field of *computational workload reduction* has emerged in recent years and numerous methods have been proposed which can help to mitigate some of the costs of the physical infrastructure. The goal of such methods is the reduction of the required amount of computations for some specific tasks in the biometric recognition pipeline. As the computational costs of the biometric template comparisons typically dominates the overall computational effort in biometric identification transactions, most of the approaches proposed in the literature are aimed specifically at optimising this step of the biometric identification pipeline [14]. More specifically, two broad classes of approaches can be distinguished: *pre-selection*, concentrating on the reduction of the search space, *i.e.* the number of necessary template comparisons (see *e.g.* [45]), and *feature transformation*, aimed at lowering the computational cost of the individual template comparisons (see *e.g.* [46]). The former are of interest in the context of this article.

Numerous methods rely on the so-called pre-filtering of the enrolment database during a biometric identification transaction. Such methods depend on categorical or weakly discriminative features (*e.g.* geographic and/or demographic metadata [47] or soft biometrics [48]), whereby the potential search space can be narrowed down quickly prior to considering the actual highly discriminative, but more computationally expensive to compute, biometric features. Conceptually similar two or multi-stage methods operating on weakly discriminative, compact representations (*e.g.* dimensionally-reduced or binarised) representation of biometric data have also been considered [49]–[51]. Likewise, general concepts of coarse-to-fine search, nearest-neighbour search, and clustering based on the feature sets extracted from biometric samples have also been proposed [14], [45].

More complex methods directly utilising the extracted biometric features and aimed at creating an intelligent search structure (*e.g.* a search tree) have been shown to be capable of significantly reducing the computational workload. In [52], a tree-based indexing and retrieval system for iris data has been proposed. Many successful methods of biometric indexing integrate information fusion; for example, [53] for multi-instance fingerprint and iris data, respectively. Furthermore, generic multi-biometric indexing methods have also been proposed *e.g.* in [54]–[56].

In [57], a multi-biometric cascade has been proposed with the aim of successively filtering the candidate short-lists based on score-level information. Similar concepts were utilised in [58], where a signal-level fusion (*i.e.* morphing, see *e.g.* [59]) of facial images facilitates a computationally efficient and accurate indexing and retrieval for biometric identification. Those methods are most closely related to the indexing and retrieval method presented in this article.

Generally, the methods mentioned in this subsection often require the storage of additional information (*e.g.* metadata) and/or a kind of a ''setup'' step (*e.g.* creation of a search structure) which requires some computational effort, but only needs to be performed infrequently. On the other hand, many of the described methods facilitate the reduction of computational workload associated with biometric identification transactions by several orders of magnitude w.r.t. the typical exhaustive-search based retrieval method.

C. BIOMETRIC TEMPLATE PROTECTION

Biometric template protection represents an active field of research since more than two decades. Comprehensive surveys on this topic can be found in [19]–[22]. Biometric template protection methods are usually categorised as *cancelable biometrics* and *biometric cryptosystems*. Cancelable biometrics employ transforms in signal or feature domain which enable a biometric comparison in the transformed (encrypted) domain [62]. Biometric cryptosystems

commonly bind a key to a biometric feature vector resulting in a protected template. Biometric comparison is then performed indirectly by verifying the correctness of a retrieved key [63].

In addition to such domain-specific approaches, general-purpose homomorphic encryption can be employed for biometric template protection [64]. Homomorphic encryption makes it possible to compute operations in the encrypted domain which are functionally equivalent to those in the plaintext domain and thus enables the estimation of certain distances between homomorphically encrypted biometric templates. As defined in ISO/IEC IS 24745 [22], biometric template protection schemes shall fulfil the following requirements:

Unlinkability the infeasibility of determining if two or more protected templates were derived from the same biometric instance, *e.g.* face. By fulfilling this property, cross-matching across different databases is prevented.

Irreversibility the infeasibility of reconstructing the original biometric data given a protected template and its corresponding auxiliary data. With this property fulfilled, the privacy of the users' data is increased, and additionally the security of the system is increased against presentation and replay attacks. Depending on the used template protection method, guaranteeing this property may rely on sufficiently protecting a certain secret (*e.g.* private encryption key(s)) from being compromised by an attacker.

Renewability the possibility of revoking old protected templates and creating new ones from the same biometric instance and/or sample, *e.g.* face image. With this property fulfilled, it is possible to revoke and reissue the templates in case the database is compromised, thereby preventing misuse.

Performance preservation the requirement of the biometric performance not being significantly impaired by the protection scheme.

Table 1 lists the mentioned types of biometric template protection and their properties w.r.t. the above criteria as well as key derivation and efficient biometric comparison. The majority of approaches on cancelable biometrics and biometric cryptosystems report a performance gap between protected and original (unprotected) systems [21], as opposed to approaches employing homomorphic encryption. Cancelable biometrics usually employ a biometric comparator similar or equal to that of unprotected biometric systems. Therefore, cancelable biometrics are expected to maintain the comparison speed of the unprotected system which makes them also suitable for biometric identification [14]. In contrast, biometric cryptosystems may need more complex comparators. Similarly, homomorphic encryption usually requires higher computational effort. Practical applications of certain template protection methods, *e.g.* homomorphic encryption, rely on maintaining the secrecy of the private key(s) used to protect the data (see also subsection V-E).

TABLE 1. Properties of template protection categories.

Template protection category	Unlinkability	Irreversibility	Renewability	Performance preservation	Efficient comparison	Key derivation
Cancelable biometrics	✓	✓	✓	(✓)	✓	
Biometric cryptosystems	✓	✓	✓	(✓)	(✓)	✓
Homomorphic encryption	✓	✓	✓	✓	(✓)	

Some research efforts and standardisation activities have been devoted to establishing metrics for evaluating the aforementioned properties of biometric template protection schemes, *e.g.* in [65]–[68]. Nonetheless, additional specific cryptanalytic methods may be necessary to precisely estimate the security/privacy protection achieved by a particular template protection scheme. Moreover, the result of such an evaluation also depends on the biometric data to which the template protection system is applied. This makes a comparison of published results difficult and sometimes misleading.

In 2001, Ratha *et al.* [69] proposed the first cancelable face recognition system using image warping to transform biometric data in the image domain. Another popular cancelable transformation of face images based on random convolution kernels was presented in [70]. In contrast to [69], this approach employs a fundamentally reversible distortion of the biometric signal based on some random seed which later coined the term ''biometric salting''. The majority of published cancelable face recognition schemes applies transformations in the feature domain [62]. Over the past years, numerous feature transformations have been proposed in order to construct face-based cancelable biometrics, *e.g.* BioHashing [71], BioTokens [72], and Bloom filters [73]. Recently, feature transformations have been specifically designed for deep convolutional neural networks, *e.g.* learned security [74]. Analyses of some popular cancelable face recognition systems have uncovered security gaps, *e.g.* in [75]–[78], and already led or are expected to lead to (continuous) improvements of such schemes.

Regarding biometric cryptosystems, the fuzzy commitment scheme [79] and the fuzzy vault scheme [80] represent widely used cryptographic primitives. Both schemes enable an error-tolerant protection of (biometric) data by binding them with a secret, *i.e.* key. Binarised face feature vectors have been protected through the fuzzy commitment scheme in various scientific publications, *e.g.* in [81], [82]. Also some works have employed the fuzzy vault scheme for face template protection, *e.g.* in [83]–[85]. It is worth mentioning that some template protection approaches combine concepts of cancelable biometrics with those of biometric cryptosystems resulting in hybrid schemes [20].

For a long time, homomorphic encryption has been considered as impractical for biometric template protection

TABLE 2. Overview of most relevant privacy-preserving WR schemes for face-based identification systems (results reported for best configurations and scenarios; note the differences in the used evaluation datasets and performance metrics).

Approach	Workload reduction category	Template protection category	Dataset	Biometric performance
Wang *et al.* [60]	Feature transformation Pre-selection	Not traditional BTP	FERET LFW	89% HR 95% HR
Murakami *et al.* [24]	Feature transformation	Cancelable biometrics	NIST BSSR1 SET3	0.1% FRR, 0.022% FAR
Dong *et al.* [23]	Feature transformation	Cancelable biometrics	LFW (closed-set) LFW (open-set) VGG2 (closed-set) VGG2 (open-set) IJB-C (closed-set) IJB-C (open-set)	99.75% RR-1 97.99% DIR, 1% FAR 99.03% RR-1 96.03% DIR, 1% FAR 80.57% RR-1 56.80% DIR, 1% FAR
Sardar *et al.* [25]	Feature transformation	Cancelable biometrics	CASIA-V5 IITK CVL FERET	99.85% CRR-1 100% CRR-1 100% CRR-1 100% CRR-1
Drozdowski *et al.* [27]	Feature transformation	Homomorphic encryption	FERET	∼5% FNIR, 1% FPIR
Engelsma *et al.* [28]	Feature transformation	Homomorphic encryption	MegaFace	81.4% RR-1
Osorio-Roig *et al.* [61]	Pre-selection	Homomorphic encryption	FEI FERET LFW	0.0% FPIR, 0.0% FNIR 0.0% FPIR, 0.2% FNIR 1.0% FPIR, 2.5% FNIR
Ours	Pre-selection	Homomorphic encryption	MORPH	99.94% RR-1, 0.1% FPIR, 0.42% FNIR

HR: Hit Rate
FRR: False Rejection Rate
FAR: False Acceptance Rate
RR-1: Rank-1 Identification Rate
DIR: Detection and Identification Rate
CRR-1: Correct Recognition Rate at Rank-1
FPIR: False Positive Identification Rate
FNIR: False Negative Identification Rate

due to its computational workload. However, in the last years, homomorphic encryption has been applied effectively to face-based verification where practical processing times could be achieved on commodity hardware [26]. Depending on the used homomorphic cryptosystem, different feature type transformations might be required.

Relevant works on biometric template protection for face-based identification systems, *i.e.* one-to-many comparisons, are shown in table 2. The biometric performance metrics are reproduced in the table exactly as in the cited papers. Note that differently named metrics often correspond to the same underlying concept and/or the authors use them in a manner divergent from that specified in ISO/IEC IS 19795-1 [29]. For example, RR-1 is the same as CRR-1, while FRR and FAR roughly correspond to FNIR and FPIR.

Some of the listed approaches are cancelable biometrics which usually retain the biometric comparator of the corresponding unprotected system. As mentioned earlier, this property makes these approaches well suited to be applied in identification mode. In addition, approaches for face identification with homomorphic encryption have been proposed, *e.g.* in [27], [28]. These works, use different concepts to maximize the efficiency of homomorphic encryption, including optimisation strategies, *e.g.* batching or dimensionality reduction. In summary, it is important to note that all published works on biometric template protection for face identification employ an exhaustive search, *i.e.* these scheme scale linearly w.r.t. to the number of protected reference face templates in the database.

As mentioned in subsection II-C, a large concern in biometric system deployments is the risk of data exposure.[1] Simultaneous efficient indexing and protection of biometric data has been proposed *e.g.* in [24], [60], [87], [88]. In [27], [28], the authors explore the use of homomorphic encryption in conjunction with biometric identification and attempt to reduce computational workload by applying a packing strategy to decrease the computation between the ciphertexts or by applying different (more computationally efficient) HE schemes. In summary, coupling biometric template protection with computational workload reduction (*i.e.* ensuring privacy and computational efficiency in addition to high biometric performance) is an insufficiently addressed topic in biometric research.

III. PROPOSED SYSTEM

The high-level, conceptual overview of the proposed system is demonstrated in figure 1 (indexing) and algorithm 1 (retrieval). The proposed system relies on creation of an efficient tree-like search structure by fusing the reference templates stored in the enrolment database. Let N be the number of subjects in the enrolment database and n_i (selected

[1]This risk is not merely hypothetical – consider real hacks such as those described in [86].

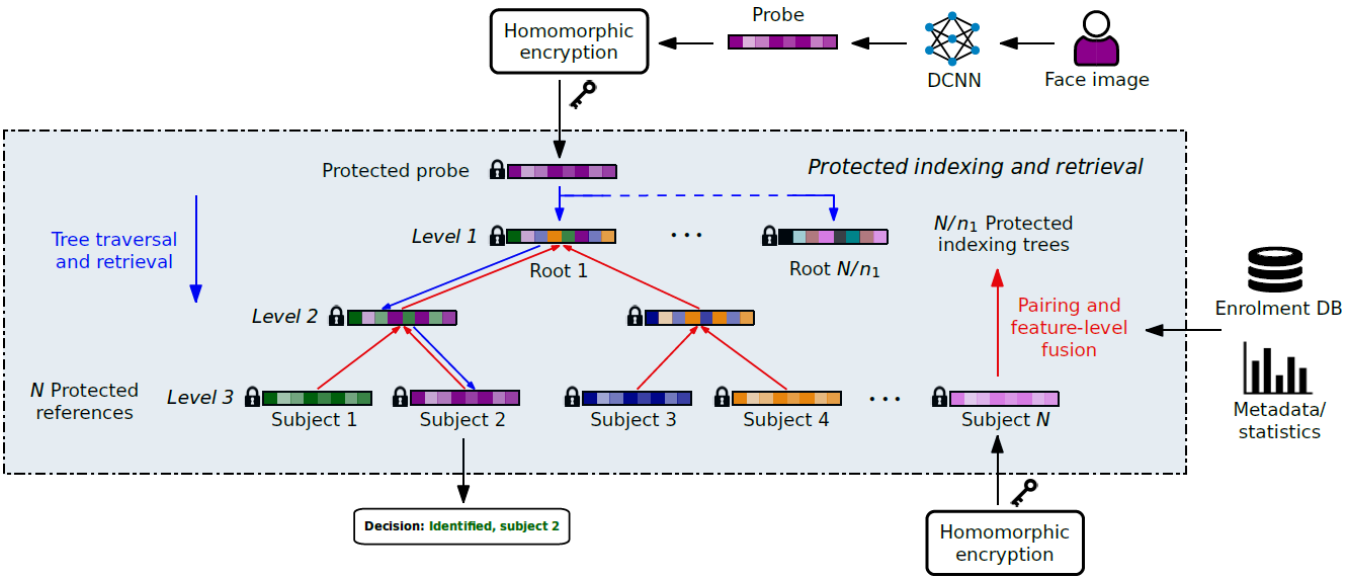


FIGURE 1. Conceptual overview of database indexing and retrieval in the proposed system.

from the set $\{2^x \mid x \in \mathbb{N}^+\}$) be the number of subjects contributing to the fused templates at the i'th level of the tree-like search structure. For instance, in figure 1, the roots of the indexing trees consist of four subjects, *i.e.* $n_1 = 4$. On the following levels, this number decreases ($n_2 = 2$) until non-fused reference templates are considered at the final level ($n_3 = 1$). Subsections III-B and III-C provide details on how the templates to be fused are paired and what information fusion methods are used.

During a biometric identification transaction, the created search structure is traversed whereby the biometric probe is compared against the fused templates in order to successively narrow down the list of potential candidate identities at each level of the search structure. The search structure has $\log_2 n_1 + 1$ levels; let k_i represent the fraction of the fused templates and their corresponding identities selected at the i'th level of the search structure. The key idea here is for k to be relatively small and decreasing at each level of the search structure. Subsection III-A provides more details on this retrieval algorithm, as well as a theoretical analysis of the possible gains in computational efficiency w.r.t. a naïve exhaustive search-based retrieval algorithm. The proposed system also allows for a seamless integration of template protection as described in subsection III-D.

A. RETRIEVAL

Since the fused templates retain sufficient discriminative power, the probes exhibit better comparison scores against their respective correct (mated) fused templates than against the other (non-mated) fused templates. Consequently, it is possible to make a robust pre-selection of a candidate short-list to be passed onto the next level of the cascade. In a successive manner, which is conceptually similar to the previous works on multi-modal and signal-level fusion-based cascades of Drozdowski *et al.* [57], [58], the candidate short-list shrinks at each level, thus resulting in fewer template comparisons being made and hence in computational workload reduction. The computational workload (W) [29] of the proposed retrieval scenario can be obtained using the following formula:

$$W = \frac{\frac{N}{n_1} + \sum_{l=2}^{\log_2 n_1} 2k_l}{N} \cdot 100\% \tag{1}$$

This equation expresses the computational workload of the proposed indexing and retrieval method as a percentage of the workload required in the typical baseline scenario where an exhaustive (1:N) search is carried out.

Figure 2 illustrates the impact of the parameters of the proposed system on its computational workload. The x-axis shows the number of fused templates at the root level of the search tree, *i.e.* how many templates are fused with each other (the n_1 parameter). The y-axis denotes the fraction of templates pre-selected at the root level (k_1) followed by a cascade with logarithmically decreasing pre-selection sizes.

Algorithm 1 Retrieval in the Proposed System

Input: *probe*, *indexing_trees*
Output: *candidates*
1: *candidates* ← roots of *indexing_trees*
2: **for** $i = 1$ to $\log_2 n_1 + 1$ **do**
3: *scores* ← compare *probe* with all *candidates*
4: *best_scores* ← find $\|scores\| \cdot k_i$ highest *scores*
5: *candidates* ← select *candidates* with *best_scores*
6: **end for**
7: **return** *candidates*

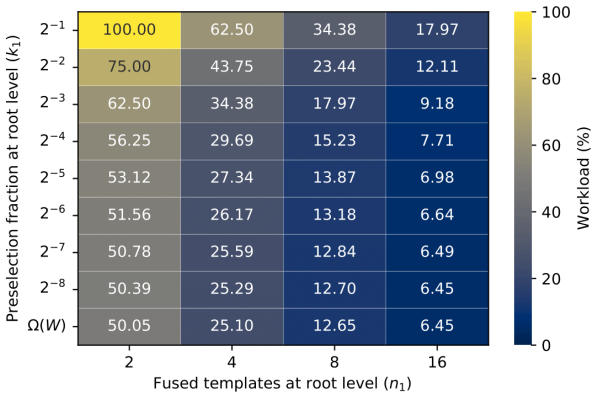


FIGURE 2. Theoretical overview of computational workload depending on the parameters of the proposed system.

Ω(W) denotes the theoretical lower bound, *i.e.* 1 fused template being pre-selected at each level of the cascade. The values in the figure are given for an N value which was used in the empirical experiments reported later on in the paper. With an increasing N value (*i.e.* growth of enrolment database size), the lower bound of computational workload for the proposed system can be expressed as follows:

$$\lim_{N\to\infty} \Omega(\mathrm{W}) = \frac{1}{n_1} \cdot 100\% \tag{2}$$

Following three observations can be made:

1) There do exist configurations (based on k_1 and n_1 parameters) which require significantly less computational workload than the baseline. In other words, provided sufficient discriminative power, the proposed system is capable of reducing the computational workload in biometric identification.
2) The k_1 parameter has a moderate impact on the possible computational workload reduction within each of the four columns. Indeed, diminishing returns are quickly approached as k_1 at root level decreases – *c.f.* the workload at lower bound with values at *e.g.* $k_1 = 2^{-4}$ or $k_1 = 2^{-5}$.
3) The n_1 parameter has a large impact on the possible computational workload reduction. Each time n_1 is doubled (*i.e.* the height of the cascade increases), the workload is approximately halved for all k_1 values.

From the above observations, it follows that for a workload-centric perspective, one would prefer as high n_1 value as possible in order to achieve highest possible workload reduction; the k_1 values usually being a secondary concern. Consider, for instance, that for $n_1 = 8$, the lower bound for achievable workload is 12.65%. If n_1 can be increased to 16, the aforementioned workload is achieved already at a relatively high fraction ($k_1 = 2^{-2}$) of templates being pre-selected at that level. It is, however, important to remember that the indexing and pre-selection may increase the false-negative errors, *i.e.* the parameters n_1 and k_1 likely cannot be set to achieve the lower bound for computational workload without simultaneously causing a significant impairment of the biometric performance. In other words, the desired reduction in computational workload needs to be feasible w.r.t. the discriminative power of the utilised recognition system. This trade-off between computational workload and biometric performance is evaluated empirically later on in this article.

B. SELECTION OF FEATURE VECTOR PAIRS

Deciding which parent samples to fuse with each other is expected to have a non-trivial impact on the efficacy of the proposed system. With an intelligent matching of the fused subject pairs, an increase in the discriminative power of the pre-selection procedure is expected, thereby improving the overall results of the proposed system in terms of biometric performance and computational workload.

Ideally, similar data subjects/samples would be fused with each other. Conceptually, matching such pairs belongs to an old and well-known class of combinatorial optimisation problems. One could formulate it in terms of a stable roommates or stable marriage problem. In practical experiments, however, such formulation has been plagued by issues related to "odd pairs" and solvability on a large set of data (see [89]–[91]). In this work, those issues are circumvented by optimising the matching algorithm with a global cost function instead of seeking a stable matching. The benefit of this approach is that some poorly matched (*i.e.* with a high cost) pairs are allowed, while the overall matchings are well-optimised for a given enrolment database. In practical experiments, this formulation (corresponding to the assignment problem) has been applied successfully.

More formally, let S represent the set of data subjects present in the enrolment database. A bijective mapping of this set to itself is sought, *i.e.* $f : S \to S$, with an additional constraint that the subjects may not be mapped to themselves, *i.e.* $\forall s \in S, f(s) \neq s$. Given a weight function $C : S \times S \to \mathbb{R}^+$, the aim of a successful mapping is to minimise $\sum_{s\in S} C_{s,f(s)}$. This work considered three methods for mapping selection:

Random samples are paired purely by chance, *i.e.* no special algorithm is used for the pair selection.

Soft-biometric similarity based on soft-biometric attributes (sex, race, age) is computed between the enrolled samples as a basis for the assignment.

Similarity-score similarity based on non-mated comparison scores between the enrolled samples computed with a facial recognition system serves as a basis for the assignment.

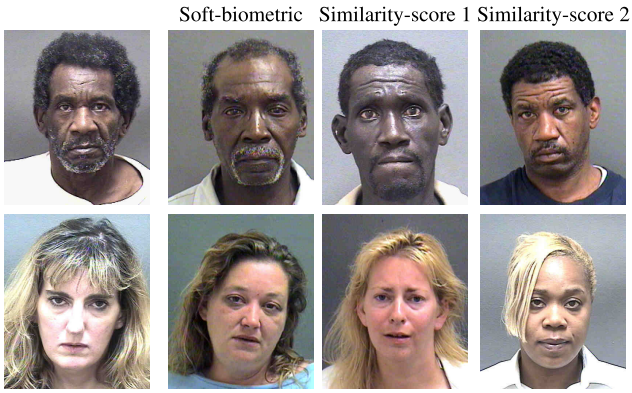
In practice, given an N-subject large enrolment database, a square matrix with the aforementioned similarity scores (soft-biometric or recognition based) can be created as illustrated in equation 3. There, S_x denotes the x'th data subject, while $c_{x,y}$ denotes the cost of pairing the x'th and y'th data subject with each other. To represent the constraint of data

subjects not being allowed to be paired with themselves, the diagonal is set to ∞. In the concrete software implementation, the largest possible value of a floating-point datatype is used instead.

$$C = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_N \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & \cdots & S_N \end{array} \\ \begin{pmatrix} \infty & c_{1,2} & c_{1,3} & \cdots & c_{1,N} \\ c_{2,1} & \infty & c_{2,3} & \cdots & c_{2,N} \\ c_{3,1} & c_{3,2} & \infty & \cdots & c_{3,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N,1} & c_{N,2} & c_{N,3} & \cdots & \infty \end{pmatrix} \end{array} \tag{3}$$

As formulated above, a *polynomial time* solution for the problem exists using the so-called Hungarian algorithm [92]. An iterative procedure is used to produce pairs for subsequent steps in the cascade, *i.e.* for $n > 2$. While nominally computationally intensive, this step is only required once (offline) during indexing of the enrolment database and not during every retrieval. Parts of the procedure can be trivially parallelised and for larger databases the polynomial factor can be mitigated by processing the database in chunks instead of directly feeding it to the algorithm.

Figure 3 shows examples of subjects paired using the soft-biometric and similarity-score based methods described above. Details on the dataset and face recognition systems used in the experiments are provided in section IV.



(a) Subject (b) Selected paired subject

FIGURE 3. Example images of pairings found using the proposed method.

C. FEATURE FUSION METHODS

The choice of information fusion method has a non-trivial impact on the discriminative power and hence the biometric performance of a recognition system [42], [43]. The overarching goal of feature-level fusion is to create a fused feature vector $\overline{\mathbf{v}} = (\overline{v}_i)_{i=1}^{n}$, $\overline{v}_i \in \mathbb{R}$ from a pair of feature vectors $\mathbf{v}$ and $\mathbf{v}'$ of same size. For simplicity, this definition and the formulas of the used fusion methods are provided in a notation for two feature vectors being fused (*i.e.* $n = 2$). However, for the specific application scenario considered in this article, they can be (and are) trivially extended to an arbitrary number (n) of feature vectors from the set $\{2^i \mid i \in \mathbb{N}^+\}$. Note that the fused feature vectors retain the length and datatype of the originally extracted feature vectors. Hence, no change of comparators is necessary to compute distances between the fused and original templates.

Three fundamentally different types of feature fusion methods are considered and described below. In the provided formulas, μ represents an overall average value of elements at a given position and is computed on a disjoint training set of feature vectors.

1) AVERAGE-BASED

An intuitive method to fuse feature vectors is averaging. Following variants of this method are considered:

Simple average The arithmetic mean of the elements at each feature position is taken: $\overline{v_i} = (v_i + v'_i)/2$.

Weighted average The arithmetic mean of the elements at each feature position is taken and additionally weighted by the distance of this element from an overall average at the given position computed on a training set: $\overline{v_i} = (v_i|v_i - \mu_i| + v'_i|v'_i - \mu_i|)/2$. In other words, elements which strongly deviate from the average are assigned more weight.

The above two methods are henceforth referred to as "Average-1" and "Average-2".

2) DISTANCE-BASED

The following methods rely on putting the values of the individual elements in relation to some overall properties of the feature vectors.

Distance from mean For each element position, the element furthest from an overall average at the given position is chosen: $\overline{v_i} = \begin{cases} v_i & |v_i - \mu_i| \geq |v'_i - \mu_i| \\ v'_i & \text{otherwise.} \end{cases}$. In other words, the element which exhibits the strongest deviation from the average at a given position is used directly.

Distance from mean rank-based We define $\#(v_i) \in \mathbb{N}$ to be the rank of v_i in the sequence of elements of $\mathbf{v}$ sorted in ascending order according to their distance to the mean μ, preserving duplicate elements. Following this operation, the element with the highest rank at a given position is chosen: $\overline{v_i} = \begin{cases} v_i & \#(v_i) \geq \#(v'_i) \\ v'_i & \text{otherwise.} \end{cases}$

The above two methods are henceforth referred to as "Distance-1" and "Distance-2".

3) INDEX-BASED

The following methods depend on the position of the elements in the feature vectors.

Section segregation A portion (*e.g.* half) of each of the contributing vectors is taken directly: $\overline{v_i} = \begin{cases} v_i & i \leq n/2 \\ v'_i & \text{otherwise.} \end{cases}$

Alternating index The feature elements are directly taken from each of the contributing vectors in an alternating manner: $\overline{v_i} = \begin{cases} v_i & i \equiv 1 \mod 2 \\ v'_i & \text{otherwise.} \end{cases}$

The above two methods are henceforth referred to as "Index-1" and "Index-2".

D. TEMPLATE PROTECTION

In the proposed scheme, template protection is facilitated through integration of homomorphic encryption. In general, an encryption algorithm E has the homomorphic property for an operation $\odot$ if it holds $E(m_1) \odot E(m_2) = E(m_1 \odot m_2), \forall m_1, m_2 \in M$, where M is the set of all possible messages. For more details on this topic, see *e.g.* a detailed survey in [93].

As shown in [27], the template comparator for biometric templates extracted from facial images can be feasibly implemented in the homomorphically encrypted domain. Depending on the format (float, integer, binary) of the feature vectors, different template comparators and encryption schemes are suitable. For the float and integer-based templates, squared Euclidean distance might be used, whereas Hamming distance is suitable for binary templates. In HE schemes which fulfil the additive and multiplicative property, the Euclidean distance can be trivially implemented to be performed in element-wise manner. If homomorphic batching is available, the element-wise distances can be computed in a vectorised operation. Furthermore, as suggested in [26], successively rotating and adding the elements of the resulting vector of element-wise distances can be used to compute the overall distance between the two feature vectors. Hamming distance, on the other hand, can be computed using an addition and a decryption with an automatic modulo-2 operation.

The key idea of using HE-based comparators is that *mathematically equivalent* operations are conducted during template comparison in the unprotected and protected domain. Thus, the protection of the templates in the proposed scheme *fundamentally* results in no loss of biometric performance (in contrast to typical biometric cryptosystems and cancelable biometrics). Using homomorphic encryption libraries described in subsection IV-C, the biometric probes, as well as the stored biometric reference templates and the index constructed from the fused templates can all be encrypted and compared in the protected domain, thereby fulfilling the biometric template protection objectives (unlinkability, irreversibility, renewability, and performance preservation) of ISO/IEC IS 24745 [22] (see subsection V-E for more details).

The proposed scheme requires a key pair to be generated and managed by the system operator. The public key is used to encrypt the enrolment database and is published so that the clients can use it to encrypt the probe templates. The system operator stores the private key and uses it to decrypt the ciphertexts containing the comparison scores, *i.e.* after the comparator has been applied in the encrypted domain. The clients do not need to manage or store additional keys, *i.e.* the aforementioned key pair is the same for all the subjects. Due to the nature of the proposed application scenario (biometric identification), creation of separate keys for each subject would result in increased complexity, as well as higher computational and storage requirements. Additional encryption (*e.g.* TLS) may be added for the transfer of data between the clients and the system operator. The presence of such keys naturally introduces the possibility of attacks aimed at guessing or theft thereof, as well as the responsibility of the system operator for the key management. Those challenges are, however, not specific to the proposed system or to homomorphic encryption in general. Any biometric system which utilises template protection, be it through dedicated biometric template protection approaches (see subsection II-C) or classic general-purpose (non-homomorphic) encryption must address the same challenges.

IV. EXPERIMENTAL SETUP

This section provides a detailed description of the setup for the conducted experiments. The used dataset and face recognition systems are described in subsections IV-A and IV-B, respectively; subsection IV-C describes the used homomorphic encryption software, while subsection IV-D gives an overview of the evaluation methodology and metrics.

A. DATASET

The academic MORPH dataset by Ricanek *et al.* [94] has been used in the experiments. A subset of images was selected based on approximate conformance with ICAO requirements for passport images [95]. As the proposed system is aimed to function with such semi-constrained images, the so-called "in-the-wild" datasets were not considered. Furthermore, the chosen dataset facilitates the soft-biometric pairing method described in subsection III-B, as groundtruth metadata is available for three demographic attributes – sex, race, and age. Figure 4 shows example images from the used dataset, while table 3 provides a numerical summary of its partitioning for the experiments.

FIGURE 4. Example images from the used dataset.

TABLE 3. Overview of the used dataset.

Partition	Subjects	Samples
Reference	4,096	4,096
Probe (enrolled)	4,096	12,939
Probe (non-enrolled)	1,935	7,123

TABLE 4. Summary of the used homomorphic encryption schemes.

Scheme	Data type	Execution time			Storage	
		Key generation	Encryption/Decryption	Comparison	Keys	Template
CKKS [98]	float	∼779 ms	∼6 ms	∼3391 ms	∼99 MB	∼516 KB
BFV [99]	integer	∼255 ms	∼76 ms	∼618 ms	∼12 MB	∼132 KB
NTRU [100]	binary	∼362 ms	∼27 ms	∼23 ms	∼1 MB	∼912 KB

B. FACE RECOGNITION SYSTEMS

Two well-known open-source face recognition systems are used in the experiments:

ArcFace A somewhat recent (initial publication in 2018), but continually improved and refined system published by Deng *et al*. [96]. The code and pre-trained model "LResNet100E-IR,ArcFace@ms1m-refine-v2" provided by the authors are used.[2]

CurricularFace A very recent system (2020) published by Huang *et al*. [97]. The code and pre-trained model "IR101" provided by the authors are used.[3]

Both systems achieve excellent biometric performance in popular large-scale face recognition benchmarks. The systems extract compact feature vectors with 512 floating-point elements. Those vectors can be seamlessly fused using the methods described in subsection III-C. Euclidean distance is used to compute the dissimilarity between two feature vectors.

C. HOMOMORPHIC ENCRYPTION

Table 4 summarises the HE schemes used to encrypt the biometric templates. Open-source implementations were used – [101] for CKKS and BFV, and [102] for NTRU. To facilitate the encryption schemes which operate using integer or binary input data, template quantisation and binarisation methods of [46] were used. The approximate execution times given in table 4 (medians over multiple runs) refer to single operations, *e.g*. a template comparison. Additionally, in section V, this benchmark is reported for an entire identification transaction in the baseline and proposed system. The timing benchmark was conducted on a freshly installed Linux Debian 10 on a commodity notebook running an Intel Core i7 2.7 GHz CPU and 16 GB DDR4 RAM. The timings were conducted in a single-threaded environment; however, it should be noted that many of the necessary computations (*e.g*. computations of comparison scores) are trivially parallelisable or distributable.

The choices of hardware and operating systems, presence of certain optimisation libraries, as well as the programming language and code quality in the actual implementations of the HE algorithms inevitably differ. The reported execution times should therefore be viewed only as a gross estimate for the computational effort required by the specific implementations of the used HE schemes on commodity hardware.

[2]https://github.com/deepinsight/insightface
[3]https://github.com/HuangYG123/CurricularFace

It should be noted that given stronger hardware (*c.f*. [28]), significantly faster execution times for the basic operations in the homomorphically encrypted domain can be achieved. Additionally, depending on the used HE scheme, it might be theoretically possible to incorporate acceleration of linear algebra operations in the encrypted domain [103], [104].

D. EVALUATION METRICS

In the experiments, the proposed method is evaluated against an exhaustive-search based baseline. Two key aspects are considered using standardised methods and metrics [29] supported by additional ones which are commonly reported in the scientific literature:

Biometric performance In closed-set identification experiments, the CMC curves, (true-positive) identification rate (IR), and rank-1 recognition rate (RR-1) are reported. In open-set identification experiments, the DET curves of false positive identification rate (FPIR) and false negative identification rate (FNIR), as well as FNIR at a decision threshold corresponding to a fixed FPIR of 0.1% (denoted $\mathbf{FNIR_{1000}}$).

Computational workload the overall computational workload (denoted W) of a single biometric identification transaction is calculated the workload reduction by the proposed scheme w.r.t. baseline is computed. This is done based on the necessary number of template comparisons and reported for the proposed system in percentage terms in relation to the exhaustive search-based baseline (*i.e*. with $W = 100\%$).

V. RESULTS

The proposed system is evaluated experimentally as follows: in subsections V-A and V-B, an analysis is conducted to establish suitable configurations which simultaneously minimise the computational workload and maximise the biometric performance. In subsection V-C, the overall results for the selected optimal configurations are reported; an ablation study is conducted in subsection V-D. Security and scalability of the proposed method are briefly discussed in subsections V-E and V-F, respectively.

A. ANALYSIS OF COMPUTATIONAL WORKLOAD

Figure 5 shows the computational workload in terms of necessary template comparisons for the proposed indexing and retrieval method. In contrast to the general, theoretical overview from subsection III-A, this figure pertains to the specific experimental setup described in section IV.

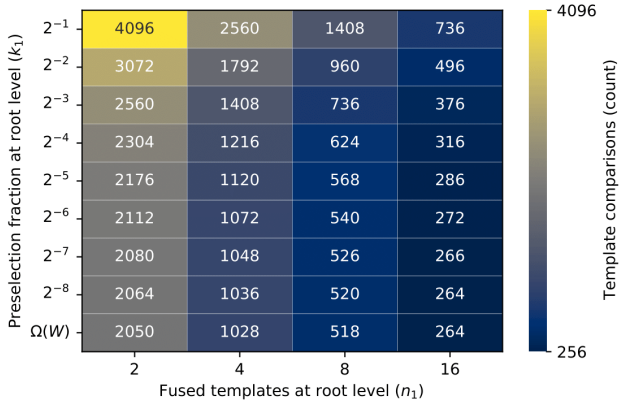


FIGURE 5. Numbers of necessary template comparisons per identification transaction for $N = 4096$ using the proposed system.

As noted in the theoretical analysis, the desired target area for the parameter (n_1 and k_1) selection lies in the bottom right corner of the matrix. For instance, given $n_1 = 16$ and $k_1 = 2^{-3}$, only 376 template comparisons are required for a biometric identification transaction, which is much lower than the 4096 template comparisons needed in the baseline scenario. In general, there exist several parameter configurations which result in the numbers of necessary template comparisons being significantly (between approximately 4 and 16 times) lower than those of a baseline retrieval method performing an exhaustive search.

In the next subsection, an analysis is conducted to determine whether the desirable configurations w.r.t. computational workload (*i.e.* the rightmost part of the matrix in figure 5) are also feasible w.r.t. biometric performance.

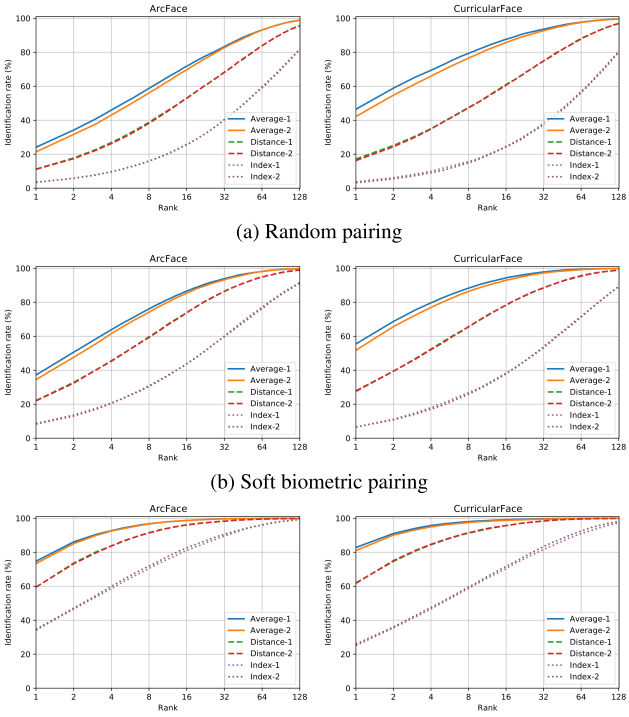
B. ANALYSIS OF PAIRING AND FUSION METHODS

From the point of view concentrated on biometric performance, the optimal selection of n_1 and k_1 parameters depends on following two factors:

1) The inherent discriminative power of the recognition system.
2) The information loss caused by the template fusion.

The information loss due to template fusion further depends on two factors: the used fusion method (section III-C) and the number of templates fused with each other (the n_1 parameter). As previously mentioned, the proposed indexing and retrieval scheme may cause false-negative errors when improperly configured, while the false-positive errors would remain unaffected or even slightly reduced through its application. To evaluate the two aforementioned factors, a closed-set identification scenario can be used and evaluated using CMC curves, which report the identification rate at given ranks (denoted r) in an ordered list of comparison scores between the probes and enrolment database.

Figure 6 shows the CMC curves for the considered facial recognition systems, template pairing methods, and template fusion methods. Aiming at highest possible workload



(a) Random pairing

(b) Soft biometric pairing

(c) Comparison score pairing

FIGURE 6. CMC curves for experiment with $n_1 = 16$ templates contributing to a fusion.

TABLE 5. Identification rates with $n_1 = 16$ (in %).

Recognition	Pairing	Fusion	Rank							
			1	2	4	8	16	32	64	128
ArcFace	Random	Average-1	24.09	34.21	45.98	58.76	71.76	83.32	93.17	98.83
		Average-2	21.38	31.69	43.01	56.12	69.59	82.61	93.08	98.75
		Distance-1	11.24	17.76	26.89	38.77	52.97	68.34	83.90	95.52
		Distance-2	11.11	17.40	26.25	38.23	52.81	68.15	83.82	95.57
		Index-1	3.40	5.85	9.62	15.93	25.78	40.18	59.64	81.66
		Index-2	3.51	5.84	9.71	15.98	25.55	40.40	59.01	81.30
	Soft biometric	Average-1	37.17	50.63	63.88	76.10	86.61	93.89	98.20	99.76
		Average-2	34.52	47.61	61.55	74.16	85.42	93.21	98.13	99.81
		Distance-1	22.15	32.86	45.23	59.64	73.98	86.43	94.85	98.95
		Distance-2	22.00	32.39	45.55	59.28	73.81	86.43	94.89	99.08
		Index-1	8.66	13.66	20.76	30.38	43.94	60.12	77.22	91.66
		Index-2	8.32	13.03	20.39	30.89	43.77	59.73	76.29	91.38
	Comparison score	Average-1	74.72	86.13	92.72	96.78	98.72	99.68	99.91	99.99
		Average-2	73.31	85.11	92.50	96.64	98.73	99.66	99.90	99.99
		Distance-1	59.43	73.65	83.72	91.37	96.00	98.35	99.54	99.89
		Distance-2	59.54	73.22	83.71	91.54	96.14	98.31	99.51	99.90
		Index-1	34.12	46.56	58.59	70.48	81.00	89.56	96.23	99.31
		Index-2	34.61	47.10	59.96	71.98	82.58	90.59	96.06	99.25
CurricularFace	Random	Average-1	46.48	58.90	69.38	79.40	87.63	93.56	97.70	99.56
		Average-2	42.17	54.76	65.95	76.61	85.71	92.70	97.53	99.54
		Distance-1	17.15	25.19	35.05	47.51	61.01	74.87	87.77	96.84
		Distance-2	16.23	24.41	34.80	47.26	60.64	74.98	87.97	96.92
		Index-1	3.79	6.16	9.95	15.72	24.69	38.24	57.02	80.64
		Index-2	3.18	5.50	9.07	14.85	24.20	37.47	56.27	79.97
	Soft biometric	Average-1	55.48	68.72	79.67	88.34	94.38	97.81	99.44	99.90
		Average-2	51.76	65.70	77.00	86.38	93.01	97.26	99.17	99.88
		Distance-1	27.52	39.36	52.48	65.48	78.26	88.54	95.37	98.92
		Distance-2	27.87	39.46	52.11	65.66	78.45	88.45	95.63	98.96
		Index-1	6.56	11.21	18.01	26.79	38.33	53.29	71.17	89.07
		Index-2	6.61	10.76	17.08	25.88	37.91	53.90	71.79	89.24
	Comparison score	Average-1	82.83	91.04	95.80	98.04	99.14	99.72	99.93	99.99
		Average-2	81.03	90.10	94.99	97.55	98.77	99.50	99.86	99.98
		Distance-1	62.04	74.63	84.43	91.34	95.72	98.41	99.61	99.93
		Distance-2	61.73	75.10	84.71	91.53	95.72	98.45	99.59	99.94
		Index-1	25.03	35.54	46.59	58.73	70.37	81.54	90.99	97.39
		Index-2	25.97	36.02	47.48	59.50	71.67	83.45	92.50	98.12

reduction (recall subsection V-A), $n_1 = 16$ (*i.e.* maximum rank is 256) is selected. Table 5 shows the numeric values of identification rate for the specific ranks depicted on the x-axis in the figure.

Following observations can be made:

1) The proposed methods of intelligent pairing of subject templates to be fused result in a significant

improvement of the identification rates w.r.t. to random pairings (*i.e.* no optimisation).

2) The pairing of templates based on (non-mated) comparison scores performs much better than the pairing based on soft-biometric attributes of the data subjects.
3) The biometric performance across the three considered types of template fusion methods varies significantly. In all considered cases, the fusion methods based on averaging perform best, relatively closely followed by fusion methods based on distance from mean. The index-based fusion methods achieve a poor biometric performance. The differences between the fusion method variants within their respective method types are insignificant.
4) Although rank-1 identification rate is very low, both recognition systems quickly converge (for the best performing type of fusion methods) at 100% well before the maximum rank of 256.
5) In general, CurricularFace performs slightly better than ArcFace. However, the differences are not very large and the general trends described above persist across both recognition systems.

Based on the above evaluations and observations, the selection of optimal configurations for pre-selection can be made. Accordingly, following choices are made:

Template pairing the method based on non-mated comparison scores is chosen.

Template fusion the method based on averaging the contributing templates is chosen.

Number of fused templates $n_1 = 16$ can be used, as both recognition systems appear to exhibit sufficient discriminative power to compensate for the information loss caused by fusing so many templates.

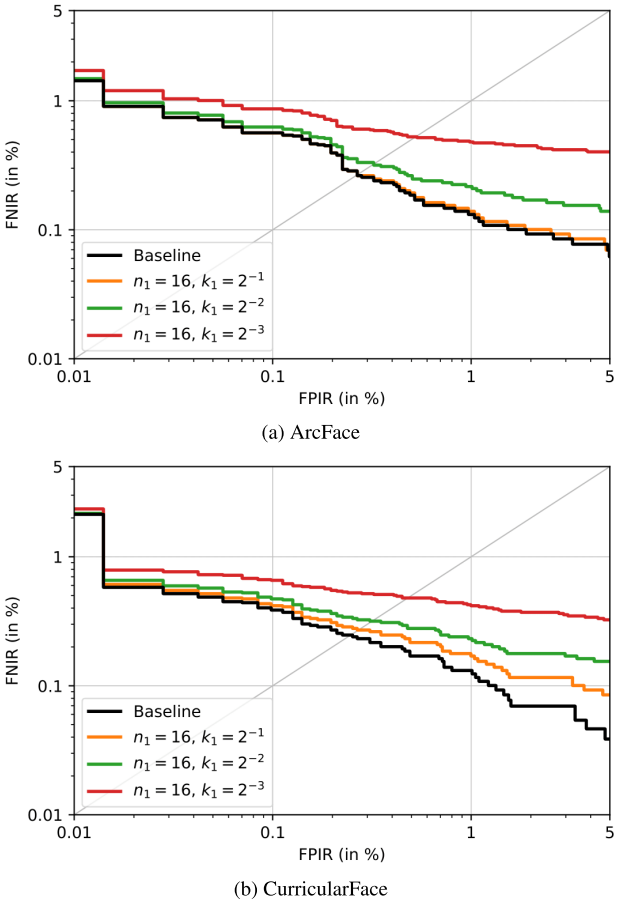
Fraction of preselected templates To avoid too many pre-selection errors, configurations with $IR(r) > 99.5\%$ are considered. This condition is satisfied for both ArcFace and CurricularFace when $r \in \{32, 64, 128\}$, using the comparison score-based pairing and averaging-based fusion. These r values correspond to $k_1 \in \{2^{-1}, 2^{-2}, 2^{-3}\}$. For recognition systems with greater discriminative power, it is conceivable to achieve even lower r and k_1 values, thereby facilitating higher workload reduction.

C. OVERALL RESULTS

To evaluate the overall performance of the proposed indexing and retrieval system, open-set and closed-set identification experiments are carried out for the configurations selected in subsections V-A and V-B. Figure 7 shows the obtained DET curves, while table 6 reports the numeric results using metrics described in subsection IV-D.

Following observations regarding computational workload and biometric performance can be made:

ArcFace All three chosen configurations perform similarly to the baseline. The most conservative one in terms of computational workload reduction, *i.e.* $k_1 = 2^{-1}$, achieves biometric performance essentially indistinguishable from that of the baseline, while simultaneously requiring only around 18% of the computational workload that the baseline requires. The computational workload can be further reduced to less than 10% of the baseline workload ($k_1 = 2^{-3}$), while retaining a reasonable (albeit slightly reduced) biometric performance w.r.t. the baseline.



(a) ArcFace

(b) CurricularFace

FIGURE 7. DET curves for the chosen parameter configurations.

CurricularFace The results mirror those of ArcFace, thus indicating a generalisability of the proposed indexing and retrieval method. The achieved computational workload reduction is identical, as same parameter configurations have been used. The biometric performance of CurricularFace is slightly better than that of ArcFace, in particular at the FNIR_{1000} operating point. The proposed system basically maintains the biometric performance of the baseline at $k_1 \in \{2^{-1}, 2^{-2}\}$ for the practically relevant FPIR values, whereas $k_1 = 2^{-3}$ yields an even lower computational workload at the cost of a slight reduction in biometric performance.

In table 7, a summary of the computational requirements for the proposed protected indexing system is given in actual runtimes and storage usage for the off-the-shelf hardware mentioned in subsection IV-C.

TABLE 6. Summary of the proposed system results using metrics corresponding to those reported in table 2. The biometric performance for the protected and unprotected versions of the system is identical.

Recognition	Configuration	Workload	Open-set			Closed-set
			EER	FPIR	FNIR	RR-1
ArcFace	Baseline	100.00%	0.27%	0.1%	0.56%	99.96%
	$n_1 = 16, k_1 = 2^{-1}$	17.97%	0.27%		0.56%	99.95%
	$n_1 = 16, k_1 = 2^{-2}$	12.11%	0.32%		0.63%	99.87%
	$n_1 = 16, k_1 = 2^{-3}$	9.18%	0.52%		0.87%	99.57%
CurricularFace	Baseline	100.00%	0.25%	0.1%	0.39%	99.98%
	$n_1 = 16, k_1 = 2^{-1}$	17.97%	0.27%		0.42%	99.94%
	$n_1 = 16, k_1 = 2^{-2}$	12.11%	0.32%		0.47%	99.86%
	$n_1 = 16, k_1 = 2^{-3}$	9.18%	0.48%		0.66%	99.64%

TABLE 7. Approximate computational and data storage requirements using the hardware specified in subsection IV-C for the unprotected system and the proposed system with $N = 4096$ and 128 bits security.

Protection	Configuration	Execution time		Storage
		DB Encryption	Identification	
None	Baseline	—	∼50 miliseconds	∼8.5 MB
	$n_1 = 16, k_1 = 2^{-1}$	—	∼9 miliseconds	∼17 MB
	$n_1 = 16, k_1 = 2^{-2}$	—	∼6 miliseconds	∼17 MB
	$n_1 = 16, k_1 = 2^{-3}$	—	∼5 miliseconds	∼17 MB
CKKS	Baseline	∼25 seconds	∼4 hours	∼2.2 GB
	$n_1 = 16, k_1 = 2^{-1}$	∼50 seconds	∼40 minutes	∼4.4 GB
	$n_1 = 16, k_1 = 2^{-2}$	∼50 seconds	∼28 minutes	∼4.4 GB
	$n_1 = 16, k_1 = 2^{-3}$	∼50 seconds	∼21 minutes	∼4.4 GB
BFV	Baseline	∼5 minutes	∼42 minutes	∼0.55 GB
	$n_1 = 16, k_1 = 2^{-1}$	∼10 minutes	∼7.5 minutes	∼1.1 GB
	$n_1 = 16, k_1 = 2^{-2}$	∼10 minutes	∼5 minutes	∼1.1 GB
	$n_1 = 16, k_1 = 2^{-3}$	∼10 minutes	∼4 minutes	∼1.1 GB
NTRU	Baseline	∼2 minutes	∼1.5 minutes	∼3.7 GB
	$n_1 = 16, k_1 = 2^{-1}$	∼4 minutes	∼17 seconds	∼7.4 GB
	$n_1 = 16, k_1 = 2^{-2}$	∼4 minutes	∼11 seconds	∼7.4 GB
	$n_1 = 16, k_1 = 2^{-3}$	∼4 minutes	∼9 seconds	∼7.4 GB

It can be observed that:

- There exist massive differences in execution time and storage space usage between the benchmarked homomorphic encryption methods. The proposed indexing and retrieval method dramatically (order of magnitude) reduces the execution times of an identification transaction w.r.t. the baseline.
- The one-time computational costs of encrypting the enrolment database and its index are negligible. Large amount of space is required for the storage of the protected templates.
- The execution times of the proposed system with BFV and especially CKKS based encryption do not suffice for real-time deployments, but could nevertheless be feasible whenever near-instantaneous system responses are not required.
- Near-realtime runtimes are achieved for the proposed system with NTRU-based encryption. This is mostly because the Hamming weight (*i.e.* the sum of the differences between individual feature vector elements) cannot be computed in the encrypted domain using this scheme. On the other hand, in BFV- and CKKS-based schemes, the analogous sum can be (and is) computed in the encrypted domain. While in principle still secure and privacy-preserving, this means that using NTRU in the proposed system introduces a further trade-off between computational requirements and potential of some information leakage.

D. ABLATION STUDY

To further analyse the impact of the individual components of the proposed system, an ablation study is conducted. Accordingly, following four systems are considered:

1) Baseline exhaustive search without template protection.
2) Baseline exhaustive search with template protection.
3) Proposed indexing and retrieval method without template protection.
4) Proposed indexing and retrieval method with template protection.

In order to reduce redundancy w.r.t. tables 6 and 7, a configuration of the proposed system which, according to the authors, offers the best trade-offs is selected for the ablation study. Accordingly, CurricularFace-based recognition is chosen, as it performs marginally better than ArcFace. The configuration $n_1 = 16, k_1 = 2^{-1}$ of the indexing and retrieval system is chosen, as its performance is nearest to that of the baseline. Finally, NTRU homomorphic encryption is chosen, as it is the fastest of the three tested ones.

Table 8 shows the summary of the results of the study. Following conclusions can be made:

Biometric performance As expected, template protection has no effect on biometric performance. This is because the same underlying mathematical operations are computed in the protected and unprotected domain. Using the proposed indexing and retrieval scheme, the biometric performance is reduced, albeit only marginally.

Execution time The inclusion of template protection significantly increases the execution time. This increase can be partially mitigated by additionally including the proposed indexing and retrieval scheme. Indexing the database requires additional computational effort and

TABLE 8. Ablation study for selected configurations from tables 6 and 7.

Setup		Biometric performance				Execution time			Workload	Storage
Retrieval	Protection	EER	FPIR	FNIR	RR-1	Indexing	DB Encryption	Identification		
Baseline	No	0.25%	0.1%	0.39%	99.98%	—	—	∼50 miliseconds	100.00%	∼8.5 MB
	Yes						∼2 minutes	∼1.5 minutes		∼3.7 GB
Proposed	No	0.27%	0.1%	0.42%	99.94%	∼10 minutes	—	∼9 miliseconds	17.97%	∼17 MB
	Yes						∼4 minutes	∼17 seconds		∼7.4 GB

time; however, it can be completed offline and needs to be done only once. Additionally, one of the main computational cost factors, *i.e.* computing the comparison scores between the templates in the enrolment database, can be trivially parallelised or distributed.

Workload Using the proposed indexing and retrieval system, the number of necessary template comparisons is reduced more than 5-fold w.r.t. to an exhaustive search baseline – irrespective of the use of template protection.

Storage Both indexing and template protection substantially increase the amount of necessary storage space. The index approximately doubles the storage space w.r.t. the baseline, whereas protected templates require an order of magnitude more space than unprotected ones.

E. SECURITY ANALYSIS

The proposed system fulfils the biometric template protection objectives specified in ISO/IEC IS 24745 [22]:

Unlinkability the mathematical operations in HE domain (and hence the distance comparators) cannot be computed over templates encrypted using two different keys, *i.e.* linking across applications is not possible by the very nature of the used HE schemes, provided that different private keys are used by those applications. Furthermore, even in the unlikely case that two applications share the same keys, the proposed schemes operate under the concepts of semantic security; specifically, a random factor is utilised in the encryption functions. Thus, encrypting an identical plaintext twice results in completely different and indistinguishable ciphertexts. Due to such guarantees from theoretical proofs of the used HE algorithms [98]–[100], empiric evaluations *e.g.* using the framework of Gomez-Barrero *et al.* [67], [73] have not been conducted. This is because it would either be impossible to compute the distances between the protected templates (different keys across applications) or the experiment would merely measure the strength of the application-specific source of randomness (same keys across applications).

Irreversibility the used HE schemes are based on ideal lattices, *i.e.* are post-quantum-secure [105]. They provide encryption with the strength of 128, 192, or 256 bits.[4] There exists a trade-off between security and computational requirements – as the encryption strength increases, so does the computational complexity. Due to the relatively low entropy of the facial representations extracted using current state-of-the-art neural networks (see below), encryption strength of 128 bits is sufficient for the proposed system.

Renewability the HE key pair can be exchanged, whereupon the biometric templates in the enrolment database and index can be re-encrypted. As previously mentioned, this would result in set of completely new ciphertexts. Ciphertexts encrypted with the old keys would no longer be accepted by the system, since computations on ciphertexts encrypted with different keys are not possible in the used HE algorithms.

Performance preservation the comparator used in the homomorphically encrypted domain is functionally identical to that of the plaintext domain, *i.e.* it yields the same comparison scores. Hence, the HE-based template protection has no impact whatsoever on the biometric performance. The biometric performance of the proposed indexing system is nearly identical to that of the baseline.

While traditional encryption schemes may provide stronger security guarantees than homomorphic encryption, this does not constitute the actual limiting factor w.r.t. facial biometrics. The entropy of facial embeddings is considered to be much lower than the aforementioned achievable cryptographic protection levels. For example, *e.g.* in [106], it has been shown that while typical facial embeddings extracted by deep neural networks consist of 512 values, their intrinsic dimensionality is much lower (more than an order of magnitude). In other words, it is more feasible (albeit still extremely difficult) to guess a sufficiently similar facial biometric template than to guess the encryption keys.

Finally, note that such attacks aimed at guessing the biometric templates and/or encryption keys (or other secrets) are not limited to the applications of homomorphic encryption for the purpose of biometric template protection. Other types of dedicated biometric template protection approaches (recall subsection II-C) as well as classic general-purpose (non-homomorphic) encryption must likewise address those challenges.

F. SCALABILITY

As the size of the enrolment database increases, following factors within the proposed system need to be considered:

[4] According to the https://homomorphicencryption.org/ standard.

Pairing although the pairing algorithm is computationally intensive, its computational costs could be easily mitigated by distributing the computations or additionally binning the enrolment database. It should also be noted that an increased size of the enrolment database would result in a larger probability of finding suitable pairings – especially for the outlier subjects (and hence an increased discriminative power of the system).

Fusion the operations for fusing the templates are implemented efficiently using vectorised operations; these computational costs are generally are negligible, *e.g.* in comparison with those required for template pairing. Furthermore, this part of the proposed system's pipeline can be trivially parallelised and/or distributed.

Encryption the computational costs of encrypting the enrolment database and index are generally very low and can additionally be trivially parallelised. The amount of RAM required to pre-load (for use during retrieval) the entire enrolment database and its index is approximately twice that of the baseline.

Retrieval the computational workload of the proposed system scales *sub-linearly* w.r.t. to the number of enrolled subjects (as opposed to a typical baseline, which typically scales linearly). Due to a flexible design of the proposed system, a dynamic adjustment (w.r.t. enrolment database size) of the decision thresholds and pre-selection subset sizes is possible. Lastly, the underlying concepts in the proposed indexing and retrieval system can be trivially distributed or parallelised.

The pairing, fusion, and encryption operations are computed infrequently and offline; they thus do not directly influence the (online) retrieval time. Considering the execution timings in table 7, it is important to note that the experiments were carried out in a single-threaded environment on an ordinary laptop. Taking advantage of parallelisation or distribution of the computations, as well as utilising more powerful hardware, these execution times could be vastly lowered (*c.f.* [28]).

VI. CONCLUSION

In this article, a method of computationally efficient indexing and retrieval of biometric data has been presented. The proposed indexing method relies on intelligent pairing of facial parent templates based on their similarity (in terms of soft biometrics or non-mated comparison scores), followed by feature-level fusion. The created search structure facilitates a multi-step biometric identification retrieval, whereby the retrieved candidate lists are successively shortened in each step of the cascade.

In a comprehensive experimental evaluation, several different pairing and fusion methods were benchmarked for the indexing step using two modern, open-source face recognition systems. Using standardised evaluation protocols and metrics, the proposed method was shown to achieve a biometric performance nearly identical to that of an exhaustive search-based baseline; simultaneously the computational workload of biometric identification transactions has been substantially reduced (down to ∼10%). In other words, by using the proposed system during biometric identification, a tenfold reduction in the required computational effort is possible with no negative impact on the biometric performance. By integrating homomorphic encryption, the proposed system achieves post-quantum-security and the biometric template protection objectives of unlinkability, irreversibility, and renewability.

In summary, the proposed system achieves a very good balance between biometric performance, computational efficiency, and privacy protection for biometric identification scenarios.

REFERENCES

[1] A. Das, C. Galdi, H. Han, R. Ramachandra, J.-L. Dugelay, and A. Dantcheva, "Recent advances in biometric technology for mobile devices," in *Proc. Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–11.

[2] (Apr. 2016). *European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom and Justice*. Eurodac Storage Capacity Increased. Accessed: Sep. 8, 2021. [Online]. Available: https://www.eulisa.europa.eu/Newsroom/News/Pages/Eurodac-storage-capacity-increased.aspx

[3] European Commission. (2018). *Smart Borders*. Accessed: Sep. 8, 2021. [Online]. Available: https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders_en

[4] Thales. (Jan. 2021). *DHS's Automated Biometric Identification System IDENT—The Heart of Biometric Visitor Identification in the USA*. Accessed: Sep. 8, 2021. [Online]. Available: https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/customer-cases/ident-automated-biometric-identification-system

[5] K. R. Moses, P. Higgins, M. McCabe, S. Probhakar, and S. Swann, *Fingerprint Sourcebook*. Washington, DC, USA: US Department of Justice, Automated Fingerprint Identification System, 2010, pp. 1–33.

[6] Federal Bureau of Investigation. (Mar. 2021). *CODIS—NDIS Statistics*. Accessed: Sep. 8, 2021. [Online]. Available: https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics

[7] Thales. (Apr. 2021). *Automated Fingerprint Identification System (AFIS)—A Short History*. Accessed: Sep. 8, 2021. [Online]. Available: https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/biometrics/afis-history

[8] Unique Identification Authority of India, "Role of biometric technology in Aadhaar enrollment," UIDAI, New Delhi, India, Tech. Rep., Jan. 2012.

[9] A. Dalwai, "Aadhaar technology and architecture: Principles, design, best practices and key lessons," Unique Identificat. Authority India, New Delhi, India, Tech. Rep., Mar. 2014.

[10] K. W. Bowyer, E. Ortiz, and A. Sgroi, "Iris recognition technology evaluated for voter registration in Somaliland," *Biometric Technol. Today*, vol. 2015, no. 2, pp. 5–8, Feb. 2015.

[11] Consortium for Elections and Political Process Strengthening, "Assessment of electoral preparations in the Democratic Republic of the Congo," CEPPS, Washington, DC, USA, Tech. Rep., May 2018.

[12] Unique Identification Authority of India. (2021). *Aadhaar Dashboard*. Accessed: Sep. 8, 2021. [Online]. Available: https://www.uidai.gov.in/aadhaar_dashboard/

[13] L. Pascu. (Mar. 2020). *Global Biometrics Market to Surpass $45B by 2024, Reports Frost & Sullivan*. [Online]. Available: https://www.biometricupdate.com/202003/global-biometrics-market-to-surpass-45b-by-2024-reports-frost-sullivan

[14] P. Drozdowski, C. Rathgeb, and C. Busch, "Computational workload in biometric identification systems: An overview," *IET Biometrics*, vol. 8, no. 6, pp. 351–368, Nov. 2019.

[15] National Institute of Standards and Technology. (Apr. 2021). *FRVT 1:N Identification*. [Online]. Available: https://pages.nist.gov/frvt/html/frvt1N.html

[16] National Institute of Standards and Technology. (2018). *Iris Exchange (IREX)*. Accessed: Sep. 8, 2021. [Online]. Available: https://www.nist.gov/programs-projects/iris-exchange-irex-overview
[17] C. Burt. (Sep. 2019). *DHS S&T Biometric Technology Rally Results Suggest Face Best for Fast Processing*. Accessed: Sep. 8, 2021. [Online]. Available: https://www.biometricupdate.com/201909/dhs-st-biometric-technology-rally-results-suggest-face-best-for-fast-processing
[18] E. Parliament, "Regulation (EU) 2016/679," *Off. J. Eur. Union*, vol. L119, pp. 1–88, Apr. 2016.
[19] A. Cavoukian and A. Stoianov, *Biometric Encryption: The New Breed of Untraceable Biometrics*. Hoboken, NJ, USA: Wiley, 2010, pp. 655–718.
[20] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," *EURASIP J. Inf. Secur.*, vol. 2011, no. 1, pp. 1–25, Dec. 2011.
[21] K. Nandakumar and A. K. Jain, "Biometric template protection: Bridging the performance gap between theory and practice," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 88–100, Sep. 2015.
[22] *Information Technology—Security Techniques—Biometric Information Protection*, Standard ISO/IEC 24745:2011, International Organization for Standardization and International Electrotechnical Committee, Jun. 2011.
[23] X. Dong, S. Kim, Z. Jin, J. Y. Hwang, S. Cho, and A. B. J. Teoh, "Open-set face identification with index-of-max hashing by learning," *Pattern Recognit.*, vol. 103, Jul. 2020, Art. no. 107277.
[24] T. Murakami, R. Fujita, T. Ohki, Y. Kaga, M. Fujio, and K. Takahashi, "Cancelable permutation-based indexing for secure and efficient biometric identification," *IEEE Access*, vol. 7, pp. 45563–45582, 2019.
[25] A. Sardar, S. Umer, C. Pero, and M. Nappi, "A novel cancelable Face-Hashing technique based on non-invertible transformation with encryption and decryption template," *IEEE Access*, vol. 8, pp. 105263–105277, 2020.
[26] V. N. Boddeti, "Secure face matching using fully homomorphic encryption," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–10.
[27] P. Drozdowski, N. Buchmann, C. Rathgeb, M. Margraf, and C. Busch, "On the application of homomorphic encryption to face identification," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2019, pp. 1–8.
[28] J. J. Engelsma, A. K. Jain, and V. Naresh Boddeti, "HERS: Homomorphically encrypted representation search," 2020, *arXiv:2003.12197*. [Online]. Available: http://arxiv.org/abs/2003.12197
[29] *Information Technology—Biometric Performance Testing and Reporting—Part 1: Principles and Framework*, Standard ISO/IEC 19795-1:2021, International Organization for Standardization and International Electrotechnical Committee, ISO/IEC JTC1 SC37 Biometrics, 2021.
[30] A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics*. USA: Springer, 2006.
[31] *Information Technology—Biometrics—Multimodal and Other Multibiometric Fusion*, Standard ISO/IEC TR 24722:2015, ISO/IEC JTC1 SC37 Biometrics, 2nd ed., Dec. 2015.
[32] Jain and Ross, "Fingerprint mosaicking," in *Proc. IIEEE Int. Conf. Acoust. Speech Signal Process.*, May 2002, pp. 4064.
[33] G. P. Kusuma and C.-S. Chua, "PCA-based image recombination for multimodal 2D+3D face recognition," *Image Vis. Comput.*, vol. 29, no. 5, pp. 306–316, Apr. 2011.
[34] V. Kanhangad, A. Kumar, and D. Zhang, "Contactless and pose invariant biometric identification using hand surface," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1415–1424, May 2011.
[35] X. Yan, W. Kang, F. Deng, and Q. Wu, "Palm vein recognition based on multi-sampling and feature-level fusion," *Neurocomputing*, vol. 151, no. 2, pp. 798–807, Mar. 2015.
[36] R. Snelick, M. Indovina, J. Yen, and A. Mink, "Multimodal biometrics: Issues in design and testing," in *Proc. 5th Int. Conf. Multimodal Interfaces (ICMI)*, 2003, pp. 68–72.
[37] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, Dec. 2005.
[38] A. Abaza and A. Ross, "Quality based rank-level fusion in multibiometric systems," in *Proc. IEEE 3rd Int. Conf. Biometrics, Theory, Appl., Syst.*, Sep. 2009, pp. 1–6.
[39] A. Kumar and S. Shekhar, "Personal identification using multibiometrics rank-level fusion," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 41, no. 5, pp. 743–752, Sep. 2011.
[40] S. Prabhakar and A. K. Jain, "Decision-level fusion in fingerprint verification," *Pattern Recognit.*, vol. 35, no. 4, pp. 861–874, Apr. 2002.
[41] P. P. Paul, M. L. Gavrilova, and R. Alhajj, "Decision fusion for multimodal biometrics using social network analysis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 11, pp. 1522–1533, Nov. 2014.
[42] L. M. Dinca and G. P. Hancke, "The fall of one, the rise of many: A survey on multi-biometric fusion methods," *IEEE Access*, vol. 5, pp. 6247–6289, 2017.
[43] M. Singh, R. Singh, and A. Ross, "A comprehensive overview of biometric fusion," *Inf. Fusion*, vol. 52, pp. 187–205, Dec. 2019.
[44] J. Daugman, "Biometric decision landscapes," Comput. Lab., Univ. Cambridge, Cambridge, U.K., Tech. Rep. UCAM-CL-TR-482, Jan. 2000.
[45] I. Kavati, M. Prasad, and C. Bhagvati, "Search space reduction in biometric databases: A review," in *Computer Vision: Concepts, Methodologies, Tools, and Applications*. Hershey, PA, USA: IGI Global, 2018, ch. 11, pp. 1600–1626.
[46] P. Drozdowski, F. Struck, C. Rathgeb, and C. Busch, "Benchmarking binarisation schemes for deep face templates," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 191–195.
[47] C. Gehrmann, M. Rodan, and N. Jönsson, "Metadata filtering for user-friendly centralized biometric authentication," *EURASIP J. Inf. Secur.*, vol. 2019, no. 1, p. 7, Jun. 2019.
[48] A. Dantcheva, P. Elia, and A. Ross, "What else does your biometric data reveal? A survey on soft biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 3, pp. 441–467, Mar. 2016.
[49] J. E. Gentile, N. Ratha, and J. Connell, "An efficient, two-stage iris recognition system," in *Proc. IEEE 3rd Int. Conf. Biometrics, Theory, Appl., Syst.*, Sep. 2009, pp. 211–215.
[50] S. Billeb, C. Rathgeb, M. Buschbeck, H. Reininger, and K. Kasper, "Efficient two-stage speaker identification based on universal background models," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2014, pp. 1–6.
[51] A. Pflug, C. Rathgeb, U. Scherhag, and C. Busch, "Binarization of spectral histogram models: An application to efficient biometric identification," in *Proc. IEEE 2nd Int. Conf. Cybern. (CYBCONF)*, Jun. 2015, pp. 501–506.
[52] P. Drozdowski, C. Rathgeb, and C. Busch, "Bloom filter-based search structures for indexing and retrieving iris-codes," *IET Biometrics*, vol. 7, no. 3, pp. 260–268, May 2018.
[53] O. N. Iloanusi, "Fusion of finger types for fingerprint indexing using minutiae quadruplets," *Pattern Recognit. Lett.*, vol. 38, pp. 8–14, Mar. 2014.
[54] U. Jayaraman, S. Prakash, and P. Gupta, "Indexing multimodal biometric databases using Kd-tree with feature level fusion," in *Proc. Int. Conf. Inf. Syst. Secur.* Berlin, Germany: Springer, 2008, pp. 221–234.
[55] A. Gyaourova and A. Ross, "A coding scheme for indexing multimodal biometric databases," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2009, pp. 93–98.
[56] A. Gyaourova and A. Ross, "Index codes for multibiometric pattern retrieval," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 518–529, Apr. 2012.
[57] P. Drozdowski, C. Rathgeb, B.-A. Mokroß, and C. Busch, "Multi-biometric identification with cascading database filtering," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 2, no. 3, pp. 210–222, Jul. 2020.
[58] P. Drozdowski, F. Stockhardt, C. Rathgeb, and C. Busch, "Signal-level fusion for indexing and retrieval of facial biometric data," 2021, *arXiv:2103.03692*. [Online]. Available: http://arxiv.org/abs/2103.03692
[59] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch, "Face recognition systems under morphing attacks: A survey," *IEEE Access*, vol. 7, pp. 23012–23026, 2019.
[60] Y. Wang, J. Wan, J. Guo, Y.-M. Cheung, and P. C. Yuen, "Inference-based similarity search in randomized Montgomery domains for privacy-preserving biometric identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 7, pp. 1611–1624, Jul. 2017.
[61] D. Osorio-Roig, C. Rathgeb, P. Drozdowski, and C. Busch, "Stable hash generation for efficient privacy-preserving face identification," *Trans. Biometrics, Behav., Identity Sci.*, Jul. 2021.
[62] V. M. Patel, N. K. Ratha, and R. Chellappa, "Cancelable biometrics: A review," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 54–65, Sep. 2015.
[63] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain, "Biometric cryptosystems: Issues and challenges," *Proc. IEEE*, vol. 92, no. 6, pp. 948–960, Jun. 2004.

[64] C. Aguilar-Melchor, S. Fau, C. Fontaine, G. Gogniat, and R. Sirdey, "Recent advances in homomorphic encryption: A possible future for signal processing in the encrypted domain," *IEEE Signal Process. Mag.*, vol. 30, no. 2, pp. 108–117, Mar. 2013.

[65] A. Nagar, K. Nandakumar, and A. K. Jain, "Biometric template transformation: A security analysis," in *Media Forensics and Security II*, vol. 7541. Bellingham, WA, USA: SPIE, 2010, pp. 237–251.

[66] Y. Wang, S. Rane, S. C. Draper, and P. Ishwar, "A theoretical analysis of authentication, privacy, and reusability across secure biometric systems," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1825–1840, Dec. 2012.

[67] 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.

[68] *Information Technology—Performance Testing of Biometric Template Protection*, Standard ISO/IEC 30136:2018, ISO/IEC JTC1 SC37 Biometrics, International Organization for Standardization, 2018.

[69] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Syst. J.*, vol. 40, no. 3, pp. 614–634, Apr. 2001.

[70] M. Savvides, B. V. K. V. Kumar, and P. K. Khosla, "Cancelable biometric filters for face recognition," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, 2004, pp. 922–925.

[71] A. B. J. Teoh, A. Goh, and D. C. L. Ngo, "Random multispace quantization as an analytic mechanism for BioHashing of biometric and random identity inputs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 1892–1901, Dec. 2006.

[72] T. Boult, "Robust distance measures for face-recognition supporting revocable biometric Tokens," in *Proc. 7th Int. Conf. Autom. Face Gesture Recognit. (FGR)*, 2006, pp. 560–566.

[73] M. Gomez-Barrero, C. Rathgeb, J. Galbally, C. Busch, and J. Fierrez, "Unlinkable and irreversible biometric template protection based on bloom filters," *Inf. Sci.*, vols. 370–371, pp. 18–32, Nov. 2016.

[74] 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. Biometrics, Behav., Identity Sci.*, vol. 3, no. 2, pp. 180–189, Apr. 2021.

[75] 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, Jul. 2006.

[76] 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, Feb. 2017.

[77] L. Ghammam, K. Karabina, P. Lacharme, and K. Thiry-Atighehchi, "A cryptanalysis of two cancelable biometric schemes based on index-of-max hashing," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2869–2880, 2020.

[78] S. Kirchgasser, A. Uhl, Y. Martinez-Diaz, and H. Mendez-Vazquez, "Is warping-based cancellable biometrics (still) sensible for face recognition?" in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Sep. 2020, pp. 1–8.

[79] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proc. 6th ACM Conf. Comput. Commun. Secur. (CCS)*, 1999, pp. 28–36.

[80] A. Juels and M. Sudan, "A fuzzy vault scheme," in *Proc. IEEE Int. Symp. Inf. Theory*, Feb. 2002, p. 408.

[81] M. van der Veen, T. Kevenaar, G.-J. Schrijen, T. H. Akkermans, and F. Zuo, "Face biometrics with renewable templates," in *Security, Steganography, and Watermarking of Multimedia Content VIII*, vol. 6072. Bellingham, WA, USA: SPIE, 2006, pp. 205–216.

[82] M. Ao and S. Z. Li, "Near infrared face based biometric key binding," in *Proc. Int. Conf. Biometrics (ICB)*, 2009, pp. 376–385.

[83] T. Frassen, X. Zhou, and C. Busch, "Fuzzy vault for 3D face recognition systems," in *Proc. Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Aug. 2008, pp. 1069–1074.

[84] Y. Wang and K. N. Plataniotis, "Fuzzy vault for face based cryptographic key generation," in *Proc. Biometrics Symp.*, Sep. 2007, pp. 1–6.

[85] C. Rathgeb, J. Merkle, J. Scholz, B. Tams, and V. Nesterowicz, "Deep face fuzzy vault: Implementation and performance," 2021, *arXiv:2102.02458*. [Online]. Available: http://arxiv.org/abs/2102.02458

[86] United States Office of Personnel Management. (Sep. 2015). *Statement by OPM Press Secretary Sam Schumach on Background Investigations Incident*. [Online]. Available: https://www.opm.gov/news/releases/2015/09/cyber-statement-923/

[87] G. Li, B. Yang, and C. Busch, "A fingerprint indexing algorithm on encrypted domain," in *Proc. Trustcom/BigDataSE/ISPA*, Aug. 2016, pp. 1030–1037.

[88] T. Murakami, T. Ohki, Y. Kaga, M. Fujio, and K. Takahashi, "Cancelable indexing based on low-rank approximation of correlation-invariant random filtering for fast and secure biometric identification," *Pattern Recognit. Lett.*, vol. 126, pp. 11–20, Sep. 2019.

[89] B. G. Pittel and R. W. Irving, "An upper bound for the solvability probability of a random stable roommates instance," *Random Struct. Algorithms*, vol. 5, no. 3, pp. 465–486, Jul. 1994.

[90] K.-S. Chung, "On the existence of stable roommate matchings," *Games Econ. Behav.*, vol. 33, no. 2, pp. 206–230, Nov. 2000.

[91] A. Röttcher, U. Scherhag, and C. Busch, "Finding the suitable doppelgänger for a face morphing attack," in *Proc. Int. Joint Conf. Biometrics (IJCB)*, Sep. 2020, pp. 1–7.

[92] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Res. Logistics Quart.*, vol. 2, nos. 1–2, pp. 83–97, Mar. 1955.

[93] A. Acar, H. Aksu, A. S. Uluagac, and M. Conti, "A survey on homomorphic encryption schemes: Theory and implementation," *ACM Comput. Surv.*, vol. 51, no. 4, pp. 79-1–79-35, 2018.

[94] K. Ricanek, Jr., and T. Tesafaye, "MORPH: A longitudinal image database of normal adult age-progression," in *Proc. 7th Int. Conf. (FGR)*, Apr. 2006, pp. 341–345.

[95] International Civil Aviation Organization, "Machine readable passports—Part 9—Deployment of biometric identification and electronic storage of data in eMRTDs," ICAO, Montreal, QC, Canada, Tech. Rep. 9303, 2015.

[96] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 4690–4699.

[97] Y. Huang, Y. Wang, Y. Tai, X. Liu, P. Shen, S. Li, J. Li, and F. Huang, "CurricularFace: Adaptive curriculum learning loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5901–5910.

[98] J. H. Cheon, A. Kim, M. Kim, and Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," in *Proc. ASIACRYPT*. Cham, Switzerland: Springer, 2017, pp. 409–437.

[99] J. Fan and F. Vercauteren, "Somewhat practical fully homomorphic encryption," *IACR Cryptol. ePrint Arch.*, vol. 2012, p. 144, Mar. 2012.

[100] D. Stehlé and R. Steinfeld, "Making *NTRU* as secure as worst-case problems over ideal lattices," in *Proc. Annu. Int. Conf. Theory Appl. Cryptograph. Techn.* Berlin, Germany: Springer, May 2011, pp. 27–47.

[101] (Nov. 2020). *Microsoft SEAL (Release 3.6)*. Microsoft Research. Redmond, WA, USA. [Online]. Available: https://github.com/Microsoft/SEAL

[102] K. Ruhloff, D. Cousins, and Y. Polyakov. (2017). *The PALISADE Lattice Cryptography Library*. Accessed: Sep. 8, 2021. [Online]. Available: https://git.njit.edu/palisade/PALISADE

[103] E. Crockett, "A low-depth homomorphic circuit for logistic regression model training," Cryptol. ePrint Arch., Tech. Rep. 2020/1483, 2020. [Online]. Available: https://eprint.iacr.org/2020/1483

[104] H. Zong, H. Huang, and S. Wang, "Secure outsourced computation of matrix determinant based on fully homomorphic encryption," *IEEE Access*, vol. 9, pp. 22651–22661, 2021.

[105] D. Bernstein and T. Lange, "Post-quantum cryptography," *Nature*, vol. 549, no. 7671, pp. 188–194, 2017.

[106] S. Gong, V. N. Boddeti, and A. K. Jain, "On the intrinsic dimensionality of image representations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3987–3996.

PAWEL DROZDOWSKI is currently a Senior Researcher with the Faculty of Computer Science, Hochschule Darmstadt (HDA), Germany. He coauthored over 20 technical publications in the field of biometrics. His research interests include biometrics, information security and privacy, pattern recognition, and algorithmic fairness. He is a member of the European Association for Biometrics (EAB) and represents the German Institute for Standardization (DIN) in ISO/IEC SC37 JTC1 SC37. He received the Best Student Paper Runner-Up (WIFS'18) Award and the Best Poster (BIOSIG'19) Award.

FABIAN STOCKHARDT is currently pursuing the M.Sc. degree with the Faculty of Computer Science, Hochschule Darmstadt (HDA), Germany. He also works as a Research Assistant with HDA. He is a member of the da/sec – Biometrics and Internet Security Research Group and the National Research Center for Applied Cybersecurity (ATHENE), Germany. He coauthored several technical publications in the field of biometrics. His research interests include image processing and biometrics, in particular face recognition.

CHRISTIAN RATHGEB is currently a Senior Researcher with the Faculty of Computer Science, Hochschule Darmstadt (HDA), Germany. He is a Principal Investigator with the National Research Center for Applied Cybersecurity (ATHENE). He coauthored over 100 technical articles in the field of biometrics. His research interests include pattern recognition, iris and face recognition, security aspects of biometric systems, secure process design, and privacy enhancing technologies for biometric systems. He is a winner of the EAB–European Biometrics Research Award 2012, the Austrian Award of Excellence 2012, Best Poster Paper Awards (IJCB 2011, IJCB 2014, and ICB 2015), and the Best Paper Award Bronze (ICB 2018). He is a member of the European Association for Biometrics (EAB), a Program Chair of the International Conference of the Biometrics Special Interest Group (BIOSIG), and an Editorial Board Member of *IET Biometrics* (IET BMT).

DAILE OSORIO-ROIG received the B.Sc. degree in computer science from the Technological University of Havana, in 2014. She is currently pursuing the Ph.D. degree with the Faculty of Computer Science, Hochschule Darmstadt (HDA), Germany. She joined the Advanced Technologies Application Center (CENATAV), Havana, Cuba, for computer science graduate training. She is a member of the da/sec – Biometrics and Internet Security Research Group and the National Research Center for Applied Cybersecurity (ATHENE), Germany. Her research interests include pattern recognition, biometrics, and machine learning, specifically, biometric indexing and privacy-enhancing technologies.

CHRISTOPH BUSCH (Member, IEEE) is currently a member of Norwegian University of Science and Technology (NTNU), Norway. He holds a joint appointment with Hochschule Darmstadt (HDA), Germany. Further he lectures biometric systems at Denmark's DTU, since 2007. On behalf of the German BSI, he has been the Co-ordinator for the project series BioIS, BioFace, BioFinger, BioKeyS Pilot-DB, KBEinweg, and NFIQ2.0. He was/is a partner of the EU projects 3D-Face, FIDELITY, TURBINE, SOTAMD, RESPECT, TReSPsS, iMARS, and others. He is also a Principal Investigator with the German National Research Center for Applied Cybersecurity (ATHENE) and is Co-Founder of the European Association for Biometrics (EAB). He coauthored more than 500 technical articles and has been a speaker at international conferences. He is a member of the editorial board of the *IET Biometrics* and formerly of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY journal. Furthermore, he chairs the TeleTrusT biometrics working group as well as the German standardization body on Biometrics and is a Convenor of WG3 in ISO/IEC JTC1 SC37.