On the Application of Homomorphic Encryption to Face Identification

P. Drozdowski*†, N. Buchmann‡, C. Rathgeb*, M. Margraf‡, and C. Busch*

* da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
† Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
‡ Freie Universität Berlin, Germany

{pawel.drozdowski,christian.rathgeb,christoph.busch}@h-da.de
{nicolas.buchmann,marian.margraf}@fu-berlin.de

***Abstract*—The data security and privacy of enrolled subjects is a critical requirement expected from biometric systems. This paper addresses said topic in facial biometric identification. In order to fulfil the properties of unlinkability, irreversibility, and renewability of the templates required for biometric template protection schemes, homomorphic encryption is utilised. In addition to achieving the aforementioned objectives, the use of homomorphic encryption ensures that the biometric performance remains completely unaffected by the template protection scheme.**

The main contributions of this paper are: It proposes an architecture of a system capable of performing biometric identification in the encrypted domain, as well as provides and evaluates an implementation using two existing homomorphic encryption schemes. Furthermore, it discusses the pertinent technical considerations and challenges in this context.

***Index Terms*—Biometric Identification, Template Protection, Face, Homomorphic Encryption**

I. Introduction

Data exposure is a potential risk in biometric system deployments, which typically store their data secured using traditional encryption algorithms [1]. If this protection were to be compromised, serious problems would arise, including but not limited to, identity theft, cross-matching without consent, and severely limited renewability. Increasingly, the public and non-governmental organisations pay attention to those (and other) issues associated with centralised storage of sensitive personal and biometric data. In some areas, this contributes to the political process of widening legislation against privacy violations (*e.g.* GDPR in Europe [2]), which entail significant responsibilities for the data controllers.

Recently, biometric *template protection* (see *e.g.* [3] for a survey) has been an active research field attempting to address said security and privacy challenges. The ISO/IEC IS 24745 [4] mandates several properties, which must be guaranteed by such schemes:

- **Unlinkability** referring to making it infeasible to determine if two or more protected templates were derived from the same instance. By fulfilling this property, cross-matching across different databases is prevented.
- **Irreversibility** referring to making it infeasible to reconstruct the original biometric data given a protected template and its corresponding secret. 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.
- **Renewability** referring to making it possible to revoke old protected templates and creating new ones from the same biometric instance and/or sample. With this property fulfilled, it is possible to revoke and reissue the templates in case of the database being compromised, thereby preventing misuse.
- **Performance preservation** referring to the requirement of the biometric performance not being significantly impaired by the protection scheme.

Three main biometric template protection approach classes can be distinguished: (1) biometric cryptosystems, which use the biometric data to bind or extract a key [5], (2) cancelable biometrics, which utilise irreversible transformations to the biometric samples or templates [6], and (3) (homomorphic) encryption of biometric data [7].

Homomorphic encryption (henceforth referred to as "HE") makes it possible to compute operations in the encrypted domain, which render the same result as those in the plaintext domain. Thus, provided that it is possible to implement a given biometric comparator to feasibly operate within the homomorphic domain, such a template protection scheme would operate without any loss of biometric performance, whereas some impairment is often inevitable in biometric cryptosystems and cancelable biometrics. 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. HE schemes are classified depending on the number and type of supported $\odot$ operations (see *e.g.* [8] for a detailed survey). The following three HE scheme types exist today:

- **Partially Homomorphic Encryption (PHE)** schemes are defined as allowing only a single operation type an unlimited number of times. PHE schemes have been around for over 30 years and are the oldest HE schemes like the classical RSA scheme [9] and the El-Gamal scheme [10]

supporting only either addition or multiplication.

- **Somewhat Homomorphic Encryption (SWHE)** schemes allow multiple operation types, but only a limited number of times. SWHE examples from literature are Yao's garbled circuit scheme [11], which supports arbitrary operations a limited number of times and the Boneh-Goh-Nissim (BGN) scheme [12], which supports unlimited number of additions and one multiplication.
- **Fully Homomorphic Encryption (FHE)** schemes support an unlimited number of operations. The first feasible FHE scheme was proposed by Gentry [13] in 2009 and many newer FHE schemes are based on Gentry's general FHE framework. Brakerski and Vaikuntanathan [14] utilise Gentry's framework to make their SWHE scheme a FHE scheme and introduce batching as an optimization [15]. The Brakerski's scheme was later optimised by Fan and Vercauteren [16].

While several authors investigated using FHE for biometric *verification* (see *e.g*. [17]–[19]) with promising results, the biometric *identification* scenario has not yet been addressed or merely considered a trivial extension of the proposed schemes. However, there exist several challenges and issues which must be dealt with for such schemes to be viable in the biometric identification scenario, especially if computational workload reduction (*i.e*. decreasing the computational complexity of the retrieval, see *e.g*. [20] for a survey) is to be employed. Accordingly, the contribution of this paper is twofold: (1) an example architecture of a system capable of performing biometric identification with homomorphically protected templates is described, implemented, and evaluated with a facial recognition system, and (2) the practical considerations and challenges relevant to the biometric identification scenario are discussed in the context of HE and potential solutions, along with the future research avenues being explored.

This paper is organised as follows: in section II the proposed system is described. Section III outlines the experimental setup and the results of the evaluations. The results and other relevant matters are discussed in section IV, while concluding remarks and a summary are given in section V.

II. Proposed System

Figure 1 shows an overview of the proposed system. There are 3 entities in the system: A client, where the biometric features are extracted (not depicted) and encrypted; a database, where the encrypted references of the enrolled subjects are stored and the distances between them and the probe computed in the encrypted domain and applies a decision threshold; and a trusted third party (TTP), which decrypts the thresholded scores and communicates a decision to the client. This description is an abstract one, the concrete HE schemes used in the implementation and evaluation are listed in section III.

The proposed system ensures that the unencrypted, privacy-sensitive biometric features are only available to the client. All network transfers of the biometric probe, as well as the computations on the data happen in the encrypted domain, thereby preventing eavesdropping attacks on the network and malicious or rogue database system attacks. Since all database entries are stored encrypted, a database breach or insider threats yield no valuable attack vectors. To further strengthen the security of the proposed system, following measures are possible: 1) to curtail active attacks (*e.g*. Man-in-the-Middle or skimming) against the transferred feature vectors and the identification transaction decision, TLS can be deployed between the parties; 2) to prevent the attacker from conducting an analysis of a single entry or unsanctioned database audits, the order of the database entries could be randomly shuffled during or after each identification transaction. Both measures would have a negligible impact on the access time and no impact on the biometric performance.

1) Let $\mathbf{v}$ denote a biometric feature vector, with a constant number (n) of feature elements: $\mathbf{v} = (v_1, v_2, \dots v_n)$ and $\mathbf{c}$ denote the element-wise encrypted version of said reference vector using a HE scheme with batching. Batching is a mechanism which allows to perform operations on vectors rather than individual numbers in the encrypted domain, thereby allowing vast (by several orders of magnitude) speed-ups on vectorisable operations. By design, the HE schemes used in this paper support a certain number of slots (s) for the elements in the encrypted vector. This number depends on the parameters of the encryption algorithm, including the encryption security. In this case (see subsection III-A), $n < s$ for any reasonable (from the security perspective) combination of parameters, which means that $\mathbf{c}$ must be padded with zeroes beyond the n-th index, *i.e*. $\mathbf{c} = (c_1, c_2, \dots c_n, 0 \dots 0)$.
2) Let $\mathbf{E}$ denote an enrolment database consisting of N such encrypted vectors (subjects), *i.e*. $\mathbf{E} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_N \end{bmatrix}$. Let $\mathbf{p}$ denote an encrypted feature vector of a probe and $\odot$ an arbitrary function for distance computation between two biometric feature vectors. In this work, the squared Euclidean distance will be used, *i.e*. $\texttt{dist}(\mathbf{c}, \mathbf{p}) = \sum_{i=0}^{s} (\mathbf{c}_i - \mathbf{p}_i)^2$. While performing the subtraction and exponentiation in the encrypted domain is done trivially by directly using the operators in the encrypted domain, the summing up of the individual elements is not as straightforward, since batching techniques do not allow access to individual encrypted vector elements. To compute the sum, the observation made in [21] is utilised – the vector is successively circularly shifted and added onto itself s times. Afterwards, the sum of elements is present in the first element of the vector. The other elements of the vector are now irrelevant and are cleaned by multiplying with 0. Mathematically, a vector $\mathbf{k} = \{1, 0, \dots 0\}$ is defined and multiplied with the result.
3) The goal of a biometric identification is to first compute

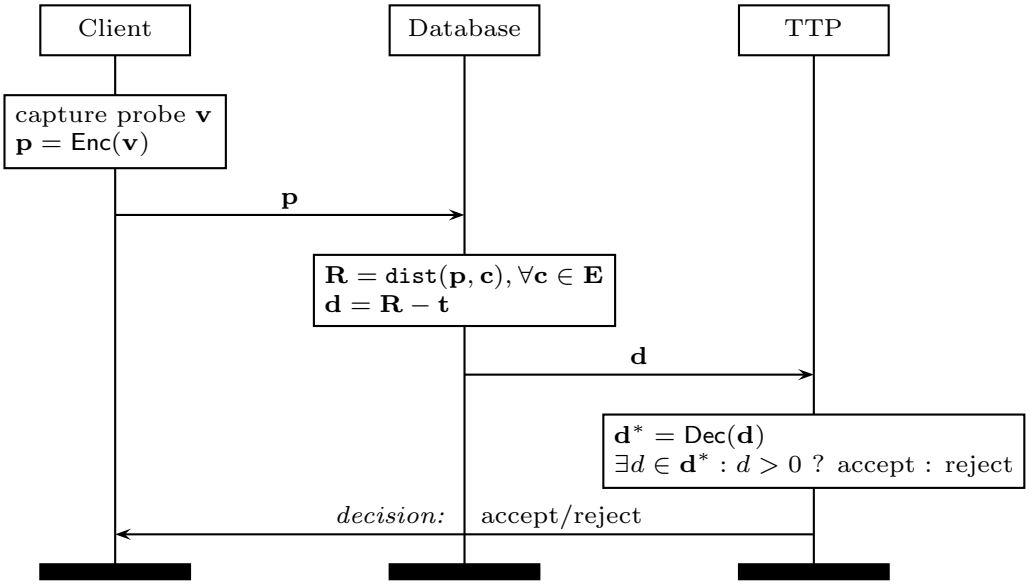

Fig. 1. System overview sketch

the comparison scores between the probe and all the items in the enrolment database, select the best one, and compare it against a predetermined threshold to make the final decision. Let $\mathbf{r}_i$ denote the result of comparison between the probe and the i'th entry in the database, *i.e.* $\mathbf{r}_i = \mathtt{dist}(\mathbf{p}, \mathbf{c}_i) \cdot \mathbf{k}$. After processing the whole enrolment database, N such result vectors as separate ciphertexts are created: $\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_N \end{bmatrix} = \begin{bmatrix} (\mathbf{r}_{1,1}, 0, \dots 0) \\ (\mathbf{r}_{2,1}, 0, \dots 0) \\ \vdots \\ (\mathbf{r}_{N,1}, 0, \dots 0) \end{bmatrix}$ note, that each of $\mathbf{r}_i$ is encrypted in a separate ciphertext, rather than $\mathbf{R}$ being encrypted in one ciphertext. In the next step, those individual ciphertexts are combined. At this point, it is also possible to randomly shuffle the order of the ciphertexts, thereby preventing the trusted third party from later learning which subject has been identified.

4) Through shifting the $\mathbf{r}_i$ vectors, a structure conceptually akin to a diagonal matrix is reached, *i.e.* $\mathbf{R} = \begin{bmatrix} (\mathbf{r}_{1,1}, 0, \dots 0) \\ (0, \mathbf{r}_{2,1}, \dots 0) \\ \vdots \\ (0, 0, \dots \mathbf{r}_{N,1}) \end{bmatrix}$. Those vectors are combined by adding them together, thereby producing a single encrypted vector holding the comparison scores of $\mathbf{p}$ against $\mathbf{E}$, *i.e.* a mapping was introduced so that $\mathbf{R} \mapsto (\mathbf{r}_{1,1}, \mathbf{r}_{2,1}, \dots \mathbf{r}_{N,1})$.
5) The next step is to compare the scores against a predefined threshold, *i.e.* to transform from the continuous spectrum of the comparison scores to the binary accept or reject decision. This is done by subtracting an encrypted vector storing at each element the threshold value (t), *i.e.* $\mathbf{t} = (t, t, \dots t)$ of length s, thus producing a decision vector $\mathbf{d} = \mathbf{R} - \mathbf{t}$, which is subsequently transmitted to the trusted third party.
6) In the last step, $\mathbf{d}$ is decrypted to determine the identification outcome (and communicate it to the client), *i.e.* decision $= \begin{cases} \text{accept} & \text{if } \exists d \in \mathbf{d} : d > 0 \\ \text{reject} & \text{if } \forall d \in \mathbf{d} : d \leq 0 \end{cases}$. If any of the elements in $\mathbf{d}$ is above 0, it is necessarily because the corresponding score in $\mathbf{R}$ has been greater than t, thus yielding an accept decision. Since due to batching access to individual elements in the encrypted vector is not possible, the whole vector must be decrypted by the trusted third party.

III. Experiments

A. Experimental Setup

The experimental evaluation was conducted on a frontal subset of the FERET database [22] in an open-set identification scenario with 500 enrolled data subjects, using 10-fold cross-validation. The features from the images were extracted using FaceNet with a pre-trained model provided by its authors [23]. FaceNet yields templates comprising of 512 floating-point feature elements; two such templates can then be compared using squared Euclidean distance. The open-source Microsoft SEAL HE library [24] was utilised to implement the identification protocol in the encrypted domain (see section II). The choice was based on the presence of a high-level API and suitable HE schemes, namely: Brakerski/Fan-Vercauteren (henceforth referred to as "BFV") [16] and Cheon-Kim-Kim-Song (henceforth referred to as "CKKS") [25] for integer and float based computations, respectively (see [8] for a detailed HE survey, including other available libraries). Thus, the templates produced by FaceNet can be encrypted directly

using CKKS. To utilise BFV, a quantisation scheme is employed, whereby the continuously distributed values of the feature elements are mapped into discrete intervals (see *e.g.* [26] for more details). Although some information is lost through quantisation, the biometric performance should remain largely unaffected. Accordingly, following evaluations were conducted on commodity hardware (one 2.5GHz CPU, 8 GB RAM) in a virtualised Linux environment:

- Biometric performance (DET curve) with the original and quantised feature vectors.
- Time elapsed (in ms) for the computations in the HE domain.

B. Results

When applying the original float-based templates and CKKS scheme, the distance computation between two encrypted feature vectors was around 5000ms. However, by applying a quantisation scheme, which enables the use of BFV, significant speed-up was achieved – the distance computation only taking 850ms. The time consumed by the encryption and decryption operations was trivial in comparison to that of the distance computation, taking around 7ms and 2.5ms for CKKS and BFV, respectively. The space requirements for the generated keys are not excessive: <1MB for the public, secret, and relinearization keys (each) and around 10MB for the Galois keys. In figure 2, it can be seen that the biometric performance of the system was not degraded by the application of quantisation.

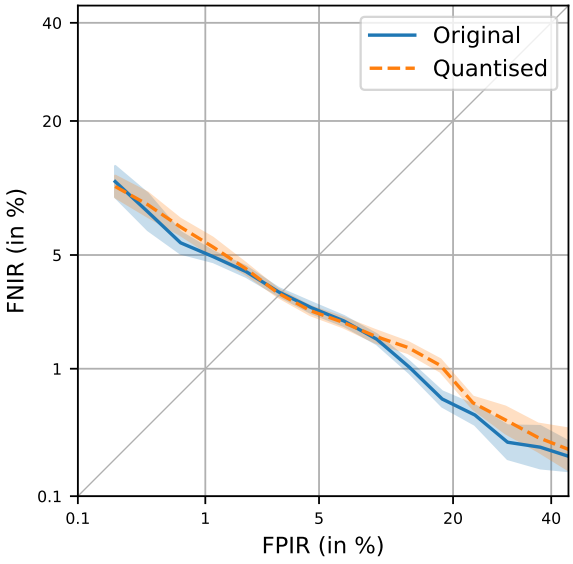

Fig. 2. DET curves for biometric identification with original and quantised features (with 95% CI)

IV. Technical Considerations

By utilising HE, the security objectives of a biometric template protection system (see section I) are achieved. The unlinkability across different databases can be guaranteed, insofar they use a different set of keys for the encryption. The irreversibility of the templates is bound to the encryption strength, which in the used library, depending on the chosen parameters, can be 128, 192, or 256 bits. Renewability is ensured, since new protected templates can always be generated by changing the encryption keys. Finally, as the template comparator in the encrypted domain is functionally identical (yields the same comparison scores) to that of the plaintext domain, the biometric performance is not impacted.

The speed of the current implementation may be prohibitive for larger deployments. It should, however, be noted that the experiments were carried out using an ordinary set-up, *i.e.* without powerful CPUs, parallelisation/distribution, *etc*. The hardware set-up notwithstanding, it would be beneficial to incorporate the concepts of computational workload reduction (see *e.g.* [20]) into such systems in order to narrow down the search space for each identification transaction. However, one drawback of using HE is that it limits the flexibility in the implementation – for instance, as previously mentioned, due to batching the feature vector elements cannot be accessed individually. This limitation makes *e.g.* the incremental recognition schemes (which facilitate early acceptance/rejection of likely/unlikely candidates) infeasible. The incorporation of more complicated schemes, such as indexing and binning, could be a potentially interesting future research avenue. On the other hand, a 1-to-first search strategy could already be implemented by slight alterations to the communication between the database and the trusted third party described in section II, however likely at the cost of at least some information exposure.

In the evaluation of biometric systems, a multitude of factors need to be considered. Some of the most important properties are the biometric performance, computational workload, as well as data security and privacy. Those goals typically counterbalance each other, and a biometric system operator is inevitably faced with trade-offs. In the case of HE in the biometric identification scenario described in this paper, currently the goals of high biometric performance, as well as data security and privacy can be fulfilled, while reducing or accelerating the computational workload could be pursued in future research to facilitate usage of such HE schemes in the practical, real-time applications.

V. Summary

In this paper, an architecture and an implementation thereof for a facial identification system in HE domain were presented and subsequently evaluated experimentally. The system fulfils the biometric template protection objectives defined in ISO/IEC IS 24745, namely unlinkability, irreversibility, renewability, and does not negatively affect the biometric performance. The paper also provided a discussion w.r.t. the technical considerations, challenges, and future work potential for utilisation of HE in the context of biometric face identification systems. While not yet practical for real-time applications, HE is definitely a promising avenue for future research and developments in this context.

ACKNOWLEDGEMENTS

This work was partially supported by the German Federal Ministry of Education and Research (BMBF), by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP), and the LOEWE-3 BioBiDa Project (594/18-17). The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the framework of MEDIAN (FKZ 13N14798).

REFERENCES

[1] K. Nandakumar and A. K. Jain, "Biometric template protection: Bridging the performance gap between theory and practice," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 88–100, September 2015.

[2] European Parliament, "Regulation (EU) 2016/679," *Official Journal of the European Union*, vol. L119, pp. 1–88, April 2016.

[3] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," *EURASIP J. on Information Security*, 2011.

[4] ISO/IEC JTC 1/SC 27 IT Security techniques, *ISO/IEC 24745:2011. Information technology – Security techniques – Biometric information protection*, International Organization for Standardization and International Electrotechnical Committee, June 2011.

[5] P. Campisi, *Security and privacy in biometrics*. Springer, 2013, vol. 24.

[6] V. M. Patel, N. K. Ratha, and R. Chellappa, "Cancelable biometrics: A review," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 54–65, 2015.

[7] 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 Processing Magazine*, vol. 30, no. 2, pp. 108–117, March 2013.

[8] 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.

[9] R. L. Rivest, A. Shamir, and L. M. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[10] T. El-Gamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *Trans. Information Theory*, vol. 31, no. 4, pp. 469–472, 1985.

[11] A. C. Yao, "Protocols for secure computations (extended abstract)," in *23rd Annual Symposium on Foundations of Computer Science*. IEEE Computer Society, 1982, pp. 160–164.

[12] D. Boneh, E.-J. Goh, and K. Nissim, "Evaluating 2-dnf formulas on ciphertexts," in *Second Theory of Cryptography Conference*, ser. Lecture Notes in Computer Science, vol. 3378. Springer, 2005, pp. 325–341.

[13] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, 2009.

[14] Z. Brakerski and V. Vaikuntanathan, "Fully homomorphic encryption from ring-LWE and security for key dependent messages," in *31st Annual Cryptology Conference*, ser. Lecture Notes in Computer Science, vol. 6841. Springer, 2011, pp. 505–524.

[15] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "(leveled) fully homomorphic encryption without bootstrapping," *TOCT*, vol. 6, no. 3, pp. 13:1–13:36, 2014. [Online]. Available: https://doi.org/10.1145/2633600

[16] J. Fan and F. Vercauteren, "Somewhat practical fully homomorphic encryption," *IACR Cryptology ePrint Archive*, vol. 2012, p. 144, 2012. [Online]. Available: http://eprint.iacr.org/2012/144

[17] M. Gomez-Barrero, E. Maiorana, J. Galbally, P. Campisi, and J. Fierrez, "Multi-biometric template protection based on homomorphic encryption," *Pattern Recognition*, vol. 67, pp. 149–163, 2017.

[18] S. Imtiyazuddin, Y. V. Subba Rao, and N. R. Rekha, "Faster biometric authentication system using Fan and Vercauteren scheme," in *Intl. Conf. on Advances in Computing, Control and Communication Technology (IAC3T)*, September 2018, pp. 48–53.

[19] V. N. Boddeti, "Secure face matching using fully homomorphic encryption," in *Intl. Conf. on Biometrics Theory, Applications and Systems (BTAS)*. IEEE, 2019, pp. 1–10.

[20] I. Kavati, M. Prasad, and C. Bhagvati, "Search space reduction in biometric databases: A review," in *Computer Vision: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018, pp. 1600–1626.

[21] C. Gentry and S. Halevi, "Implementing Gentry's fully-homomorphic encryption scheme," in *Annual intl. conf. on the theory and applications of cryptographic techniques*. Springer, 2011, pp. 129–148.

[22] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[23] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Conf. on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2015, pp. 815–823.

[24] "Microsoft seal (release 3.2)," https://github.com/Microsoft/SEAL, February 2019, microsoft Research, Redmond, WA.

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

[26] P. Drozdowski, F. Struck, C. Rathgeb, and C. Busch, "Benchmarking binarisation schemes for deep face templates," in *Intl. Conf. on Image Processing (ICIP)*. IEEE, 2018, pp. 1–5.