

# Encrypted Biometric Search: A Deep Learning Approach to Scalable and Secure Cross-Border Data Exchange

Kyriaki Miniadou

*Institute of Computer Science (ICS)*

*Foundation for Research and Technology - Hellas (FORTH)*

Heraklion, Greece

kminiadou@ics.forth.gr

Asterios Leonidis

*Institute of Computer Science (ICS)*

*Foundation for Research and Technology - Hellas (FORTH)*

Heraklion, Greece

leonidis@ics.forth.gr

Georgios Th. Papadopoulos

*Institute of Computer Science (ICS)*

*Foundation for Research and Technology - Hellas (FORTH)*

Heraklion, Greece

and

*Department of Informatics and Telematics*

*Harokopio University of Athens*

Athens, Greece

g.th.papadopoulos@hua.gr

Constantine Stephanidis

*Institute of Computer Science (ICS)*

*Foundation for Research and Technology - Hellas (FORTH)*

Heraklion, Greece

**Abstract**—Cross-border collaboration among Law Enforcement Agencies is essential for effective and timely suspect identification, especially when the availability of biometric data varies between agencies. This paper presents a scalable and secure approach for multimodal biometric identification across multiple jurisdictions. Our approach allows Law Enforcement Agencies to combine biometric modalities -facial images, fingerprints, and voice samples- and compare them with collaborating agencies, improving the overall accuracy and effectiveness of suspect identification. By leveraging deep learning models for indexing and comparison, efficient data retrieval was achieved without compromising privacy or security. To ensure the protection of sensitive biometric data, our approach incorporates advanced encryption mechanisms, including Homomorphic encryption for secure computations and AES encryption for safeguarding biometric information. Its decentralised architecture allows each LEA to maintain independent instances of the Deep Learning Indexer and Comparator, minimising risks associated with centralising sensitive data and supporting seamless collaboration between agencies. This approach not only improves the accuracy of suspect identification but also enhances operational efficiency by allowing LEAs to query and share biometric data securely across borders.

**Index Terms**—Biometrics, Multimodal Biometric Data, Deep Learning, Security, Cross-Border Collaboration

## I. INTRODUCTION

Law Enforcement operations, tasked with enforcing the law through the investigation, deterrence, or punishment of individuals who violate rules and norms governing that society, are inherently complex and sensitive. The challenges arise from legal and jurisdictional constraints -especially when national

security is at stake- and due to the management of sensitive extensive suspect information.

Efforts to foster cross-border collaboration among Law Enforcement Agencies (LEAs) have seen some initial progress [1]. However, legal complexities and data sensitivity issues are significantly amplified when operations span multiple jurisdictions [2]. Consequently, it has become imperative to develop or adapt existing frameworks utilised by a singular LEA, to ensure secure, efficient, and robust cross-border cooperation.

In this context, we propose an approach that facilitates scalable and secure cross-border multimodal identification of potential unknown suspects. This approach integrates three distinct biometric data modalities -facial images, fingerprints, and voice samples- allowing for the use of either a single modality or a combination of them to search for matches. The key advantage of cross-border matching lies in addressing situations where a LEA may not have any matching data in their databases, thus lacking critical information needed for suspect identification. Collaboration between LEAs provides a second opportunity, where even if the requesting LEA has no relevant data, the suspect may have been identified by another LEA. This approach allows requesting LEAs to search across databases of collaborating LEAs, thereby improving both identification capabilities and overall operational efficiency.

To ensure fast and efficient storage, retrieval, and comparison of inherently high-dimensional biometric data, indexing is employed. In this context hashing [3] is particularly beneficial since it converts high-dimensional biometric data into more efficient and compact representations. With the rise of deep

learning, Deep Hashing [4] has gained prominence for its ability to optimise the grouping and differentiation of similar data. This proves particularly advantageous when handling noisy biometric data that exhibit subtle variations, such as slight changes in facial expressions, fingerprint positioning or tonal differences in voice samples.

This work builds upon a previous publication on Cross-Border Collaboration by the authors [5], specifically extending the focus from Facial Biometric Search by incorporating two additional biometric modalities. It also addresses the challenges associated with combining multiple biometric modalities, as relying on a single modality may yield incomplete results, while integrating multiple modalities may sometimes lead to conflicting matching outcomes. The proposed extended approach continues to employ distributed deployment and robust encryption measures to ensure data security and privacy.

At its core the approach consists of three distinct phases -Model Training Phase, Indexing Phase and Searching Phase- each of which must be completed for the subsequent phase to be able to function properly. The Model Training Phase is responsible for training the models tailored for each modality that both the indexing and searching phases will utilise. The trained models will be distributively deployed with each LEA receiving its own snapshot of the Deep Learning Indexer and Comparator modules.

The Indexing Phase utilises the DL Indexer to process all available biometric data creating an index catalogue that stores representations of the actual data, along with a pseudonymised identifier. For each modality a different index catalogue is created and it is not imperative to acquire all modalities for a suspect. In the Searching Phase, the DL Comparator receives query biometric data and compares them against all artifacts in each index catalogue. Each biometric modality is compared exclusively against the same data type. The output is a ranked list of suspect IDs, ordered by descending similarity, which can then be used to retrieve more detailed information about the individual associated with the biometric data.

To support distributed deployment, each LEA maintains its own snapshot of the DL Indexer and DL Comparator. Additionally, all data are encrypted, using either homomorphic [6] or AES [7] encryption depending on the nature of the operations required. Homomorphic encryption is applied when computations need to be performed on the data, while AES encryption is used for biometric data containing sensitive suspect information to ensure privacy and security.

This paper is structured as follows: In Section II, prior works related to multimodal biometric frameworks are discussed. Section III, provides the motivation behind our approach, Section IV offers a detailed description of our approach, which encompasses its overall architecture. Finally, we conclude this paper in Section V.

## II. RELATED WORK

In an effort to enhance the effectiveness of integrating multiple biometric data modalities, a substantial amount of highly relevant data with superior information quality is essential.

Information fusion techniques [8] aim to improve information quality by merging heterogeneous data sources. Information quality is often compromised due to imperfections in the data or limited resources. Thus, incorporating information fusion techniques into biometric systems is vital for improving the overall performance and reliability of suspect identification efforts across jurisdictions.

In the context of biometric data evaluation, similarity can be approached in two distinct ways. Many methods emphasise a binary similar-dissimilar format, which applies a strict degree of similarity between data samples [9]. In contrast other approaches focus on similarity ranking, where a pool of existing samples is compared against a specific instance, and results are ordered in descending similarity [10]. This ranking method allows for more nuanced comparisons, offering a graded view of similarity that is particularly useful in complex biometric matching scenarios.

Regarding distributed biometric systems with enhanced security measures, an authentication framework [11] incorporated facial, fingerprint and IRIS data to enhance security and data integrity. Unlike traditional systems that rely on static parameters, this framework supported dynamic and multi-user authentication, which allows for continuous verification of user integrity in multi-user settings.

Similarly, a multimodal biometric system that combined face and fingerprint recognition was introduced in [12]. This fusion addressed the limitations of unimodal systems, such as vulnerability to noise, forgery, and a lack of universality. This approach demonstrated the potential for multimodal biometric systems to significantly enhance identification accuracy, especially in cyber-physical environments where security is paramount.

Furthermore, [13] presented a highly accurate multimodal biometric system that combined fingerprint, finger-vein, and face images using Convolutional Neural Networks for feature extraction and recognition. The system leveraged CNN-based feature extraction techniques to process each modality independently, generating feature vectors for robust person identification. The fusion of these modalities through weighted sum and product techniques further enhanced recognition accuracy.

The literature reveals a significant gap in the application of multimodal biometric approaches specifically designed to support Law Enforcement Agencies. Most approaches concern unimodal biometrics like in [14] where various algorithms for fingerprint classification and identification are examined, demonstrating their utility in criminal investigations.

For facial data [15], highlighted the challenges of using weak biometric modalities, such as facial recognition, in forensic science due to their lower discriminatory power compared to stronger biometrics like fingerprints and DNA. The proposed framework integrated these weak biometrics into forensic investigations, emphasizing semi-automatic systems where human experts review automated results to improve accuracy. Additionally, one significant project is the Speaker Identification Integrated Project (SiIP) [16], which has cre-

ated an interoperable database for voice biometrics—the third largest biometric database at Interpol after fingerprinting and facial recognition. SiiP uniquely incorporated soft biometrics like age, accent, and gender, inferred from voice data, to enhance the accuracy of speaker identification.

### III. MOTIVATION

The motivation for this work arises from the limitations of relying on a single modality of biometric data for suspect identification. In real-world scenarios, the availability of biometric data can vary significantly. For instance, in one case, a suspect may not leave fingerprints at a crime scene, but CCTV cameras may capture their facial features. In another case, the scene may be filled with the suspect’s fingerprints as they failed to conceal their identity, but no cameras were present to capture the individual. Furthermore, there are situations where CCTV footage may be insufficient due to poor lighting, yet clear audio recordings of the suspect’s voice are available. Such real-world situations are often far more complex than the controlled, hypothetical scenarios used for testing. In light of this complexity, three distinct biometric modalities—facial images, fingerprints, and voice samples—were selected to evaluate the effectiveness of this approach.

In a cross-border scenario, the most critical functionality for suspect identification is the secure transmission of suspect data between collaborating LEAs. However, this process is not as simple as sending the biometric data over the network, as numerous security and legal constraints prevent raw data from being transmitted in its original form. While a straightforward solution would involve encrypting the data before transmission and decrypting it upon arrival at the other LEA’s premises, this method introduces delays, and decrypted raw data remains vulnerable. To address this, our approach utilises Homomorphic Encryption, a form of encryption that allows computations to be performed on encrypted data without needing to decrypt it first. Since the objective is to provide potential suspect identifications ranked by similarity from a pool of existing samples in LEAs’ databases, rather than a binary similar-dissimilar sample comparison, homomorphic encryption ensures that this relative order of similarity is preserved. In essence the output of the comparison process remains the same both in encrypted and raw data computations.

Furthermore, the decision to adopt a distributed architecture in our approach stems from the limitation of a centralised solution. A centralised approach would require the encryption of all biometric data from each LEA’s databases, transmission to a central server for decryption and indexing, and then re-encryption and distribution of the indexed representations back to the respective LEAs. This process would not only be time-consuming and costly but will also pose security risks associated with the transmission of sensitive data. Additionally, when new LEAs join the collaboration, the entire databases would need to be re-indexed, further complicating the process. This distributed nature allows each LEA to maintain its own

instances of the indexer and comparator, minimising data transfer.

The innovation of the proposed approach lies not only in its ability to handle large-scale, multimodal biometric data in real-time, but also in its distributed architecture and security mechanisms, which together make an effective solution for complex applications such as cross-agency biometric matching. This distributed nature ensures that data from various sources, can be processed efficiently without centralising sensitive information such as suspect personal identification and biometric data. Moreover, the incorporation of homomorphic encryption enables the secure sharing and comparison of biometric data across agencies. This ensures that while data is processed and compared by external LEAs, it remains encrypted, protecting sensitive information in cross-agency collaborations.

### IV. THE APPROACH

The functionality of our approach relies on two key modules that leverage deep learning: the Deep Learning Indexer and the Deep Learning Comparator. Deep Learning models typically require a substantial amount of data for training to achieve robust and accurate results. However, due to the sensitive nature of biometric data, there are ethical and legal challenges in collecting real-world data. To address this, a hybrid approach was adopted. The vast majority of the data originate from publicly available and synthetic datasets that are not associated with criminal records.

Specifically, the facial indexing utilised the “FaceScrub” [17] and “YouTube Faces” [18] datasets. The former contains facial images of individuals from diverse racial and ethnic backgrounds, while the latter frames from YouTube videos uploaded by content creators. For voice data, the “CSTR’s VCTK Corpus (Centre for Speech Technology Voice Cloning Toolkit)” [19] was used, comprising voice recordings from English speaking individuals with various accents. Additionally, training for fingerprint indexing was conducted using the “FVC2000” [20] dataset, which contains fingerprint data from various participants.

For pilot testing, consortium partners have already collecting data from real users, which will be used for Machine Learning training and system demonstrations. However, only about 5% of training data will originate from datasets collected through the pilots, ensuring realistic demonstrations while maintaining the integrity of the experimental setup.

As mentioned above the functionality of this approach revolves around three distinct and sequential phases, where each phase builds upon the previous one. This cumulative structure requires that each preceding phase be completed successfully to provide the foundation for the subsequent phases to function effectively.

#### A. Model Training Phase

The Model Training Phase, as depicted in Figure 1, is responsible for developing three specialised deep learning models that are distributed to each Law Enforcement Agency in the form of a trained Deep Learning Indexer and Deep

Learning Comparator. The use of separate models for each biometric modality is critical, as deep learning models require substantial volumes of relevant data to accurately learn intricate patterns that enable accurate predictions. These specialised models are more beneficial than attempting to train a single general model that covers the prediction across all data modalities simultaneously. Furthermore, this decentralised approach is scalable, allowing for the seamless integration of new biometric data types in the future, without requiring significant modifications.

After concluding an extensive literature review and performing experiments to validate the accuracy of the models when trained with the open, publicly available datasets, two model architectures were selected. For facial data images, the Orthonormal product quantization network (OPQN) model introduced in [21] was chosen, while for fingerprints and voice samples the Deep Hashing Network (DHN) model introduced in [22] was utilised. A detailed explanation of the OPQN model's application can be found in our previous work [5]. The OPQN model creates indices denoting the closer centroid to the image and the similarity relies on distances from this predefined centroids. The DHN model generates binary hashcodes, and the comparisons between biometric data representations rely to Hamming Distance to assess similarity.

The Training Module is responsible for utilising the annotated, high-quality datasets and the identified deep learning models to produce trained model Weights. The trained weights are then distributed to each LEA by the Distributed Deployment Manager and are used to generate hash representations of the biometric data in the LEAs' databases. This decentralised approach enables the integration of additional LEAs without disrupting ongoing operations. This eliminates the need to transfer data to a central space for processing, thus enhancing security and efficiency.

Upon the completion of training, all original training data are securely deleted, and the models' outputs are designed in such a way that they cannot be used to reverse-engineer or retrieve the original biometric data. This ensures that the system remains secure, even in the unlikely event that the model weights or outputs are accessed by unauthorized entities.

### B. Indexing Phase

Once the Model Training Phase is completed, the Indexing Phase depicted in Figure 2 commences, laying the foundation for generating indexing artifacts essential for the biometric matching. The Deep Learning Indexer is designed to parse biometric data from Law Enforcement Agencies and generate an index catalogue comprising three distinct collections, each corresponding to a specific biometric modality.

Each entry in these collections is converted into a unique triplet format that contains the following information:

- Hash representation: A memory-efficient, encrypted representation of the biometric data.

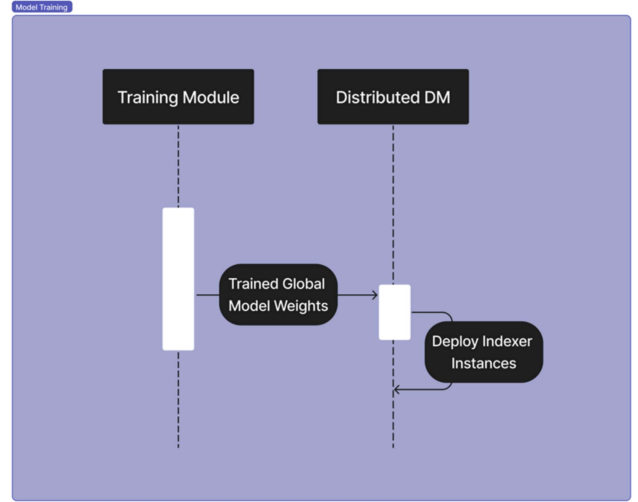


Fig. 1: Model Training Phase

- LEA ID: An unique identifier representing the originating LEA, facilitating data aggregation across multiple agencies.
- suspect ID: A pseudonymised identifier that remains pseudonymised within the distributed dataset but can be securely mapped back to real individuals by each LEA using a Pseudonymization Mapping.

The hash representation varies depending on the biometric data modality. For fingerprints and voice samples, the hashes are binary representation of the respective data. In contrast, for facial images, the hash consists of six indices, with each index representing the ID of the closest centroid to a specific part of the image. These centroids are determined by the model's weights, which are consistent in all DL-Indexers, serving as a common reference point for comparison.

To ensure secure computational operations in subsequent stages, different encryption techniques are applied. For storing the hash representations efficiently, binary hashcodes are encrypted using homomorphic encryption to allow for computational comparison operation in the later stage, while for facial images since the indexes cannot have altered values AES encryption is employed and the decryption of this indexes is needed in the subsequent phase. All generated triplets, along with the Pseudonymization Mapping, are securely stored in the Local Storage within each LEA. This storage solution provides a structured and secure repository for critical suspect information, ensuring organized access and privacy.

This indexing process allows our approach to operate in a decentralised manner, reducing risks associated with centralizing sensitive data and enabling secure collaboration between LEAs. The structured triplet format, combined with robust encryption methods, ensures that biometric data remain protected, while still enabling efficient and accurate cross-agency comparisons.

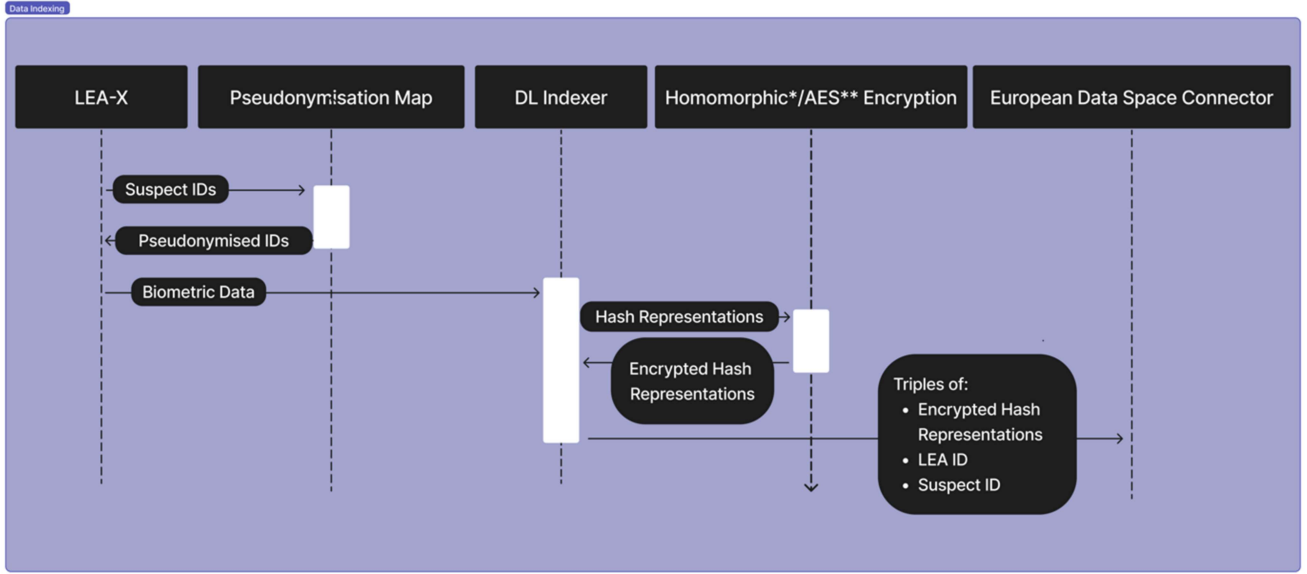


Fig. 2: Data Indexing Phase

### C. Searching Phase

Once the indexing artifacts are created, the Searching Phase, as shown in Figure 3, enables Law Enforcement Agencies to collaborate effectively in suspect identification across various jurisdictions. Each LEA maintains a local snapshot of the DL Comparator, which allows them to initiate the comparison process.

The Local DL Comparator processes biometric queries, that may include facial images, fingerprints and voice samples. Similar to the DL Indexer, the Local DL Comparator employs two distinct strategies based on the biometric modality. Fingerprints and voice samples are converted into binary representations, while facial images are transformed into an array of partial distances that representing the distance from each centroid, which is predetermined by the model's weights. This transformation is crucial for ensuring accurate biometric data comparisons. Both types of representations are then encrypted using homomorphic encryption, which preserves the relative ranking of numerical values throughout the computational process, eliminating the necessity to decrypt the query data at any stage. The order of the encrypted values mirrors that of their unencrypted counterparts. The encrypted data are transmitted to the Global DL Comparator, which acts as a broker, distributing the encrypted queries to locally deployed instances of DL Comparators belonging to collaborating LEAs.

Upon receiving homomorphically encrypted queries, each Local DL Comparator accesses the artifact collections maintained by the LEA. The Local DL Comparator retrieves the hash representations from the Local Datastore and employs two distinct comparison strategies depending on the biometric modality. For fingerprints and voice samples, Hamming Distance is employed, while for facial images, the hash representations functions as an array of indices to extract partial

distances from the query array. Rather than directly comparing hash values, the system calculates the distance between the query and the closest centroid to the hash representation. By precomputing and storing these partial distances in the query array, the method prioritises computational efficiency, though it may slightly compromise accuracy in distance calculations. The Local DL Comparator performs the comparison and similarity scoring directly on the encrypted data, ensuring that sensitive biometric information remains protected throughout the process. This mechanism preserves data privacy while allowing for accurate similarity scoring and comparison.

The Global DL Comparator upon receiving responses from all distributed instances, compiles a list of potential matches categorised by the respective authority, fostering cross-authority collaboration. The aggregated results are provided to the initiating LEA.

Each comparison result includes the following information:

- **Similarity Rate:** Ranked on a scale from *Highly Likely Suspect*, *Probable Suspect*, *Unlikely Suspect*, *Doubtful Suspect*, to *Highly Doubtful Suspect*.
- **LEA ID:** A unique identifier representing the originating LEA.
- **Suspect ID:** A pseudonymised identifier, ensuring anonymity for the suspect.

This comparison process enables our approach to function in a decentralised manner, minimising risks associated with centralised processing of sensitive biometric data and fostering secure collaboration between LEAs. By utilising encrypted query data and maintaining the integrity of similarity rankings, the DL Comparator ensures that biometric information remains protected throughout the matching process, while still providing efficient and accurate cross-agency comparisons.

To provide an intuitive visualisation for biometric data

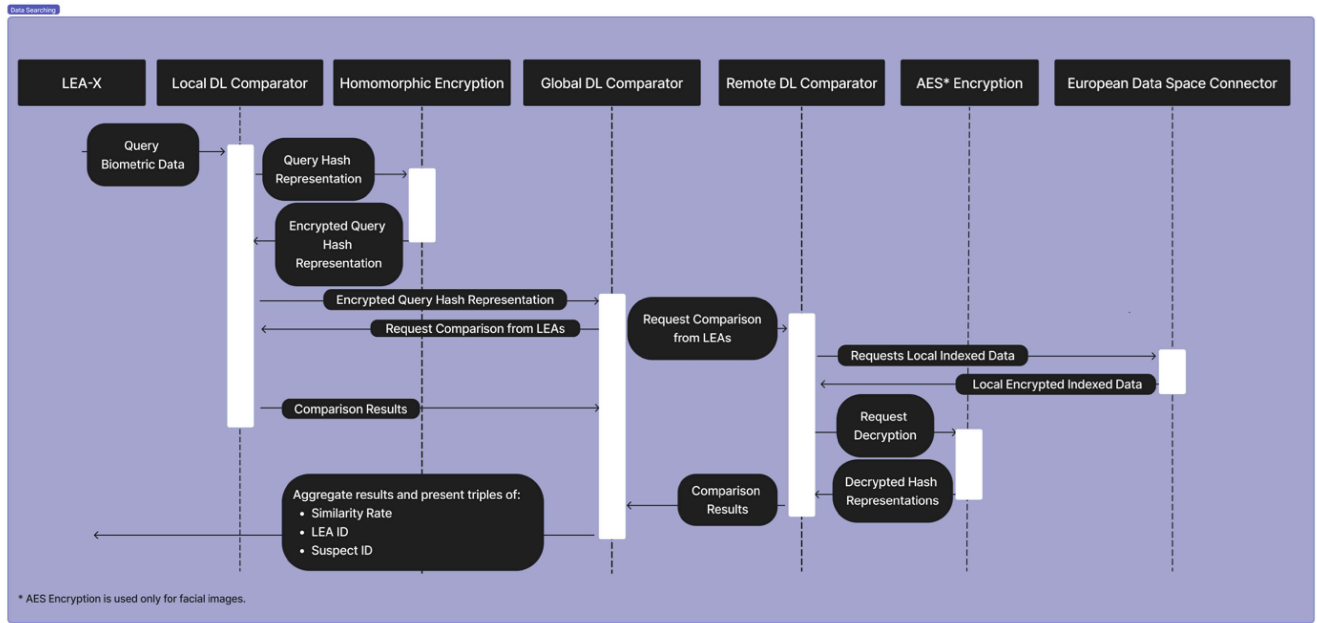


Fig. 3: Data Searching Phase

matching across multiple LEAs, as well as to avoid overwhelming law enforcement agents by providing the triples, a simple proposed visualisation is depicted in Figure 4. The results are sorted by each LEA, with the initiating LEA (PJ) shown at the top, followed by other LEAs like MOI and GPI. The information is grouped by suspect ID, and confidence scores for face, fingerprint, and voice matches are displayed if data is available. When other LEAs have suspect data but have not identified the individual, no further action can be taken. However, when an LEA has identified the individual, the system allows the user to request and access that information, supporting the investigation process. This design prioritises clarity, organising matches by confidence score and enabling efficient inter-agency collaboration while supporting actionable steps through a streamlined request system.

## V. CONCLUSION

In conclusion, this work introduces an innovative solution for cross-border biometric identification, addressing the challenges of handling multimodal data in a secure and decentralised manner. By leveraging deep learning techniques and robust encryption methods, our approach enables Law Enforcement Agencies to collaborate effectively while maintaining data privacy. The system's decentralised architecture reduces the risks associated with centralising sensitive information and facilitates efficient cross-agency suspect identification. This solution represents a significant step forward in improving public safety and operational efficiency in a global context. Future work will focus on expanding the system's capabilities and exploring additional biometric modalities for even greater accuracy and robustness.

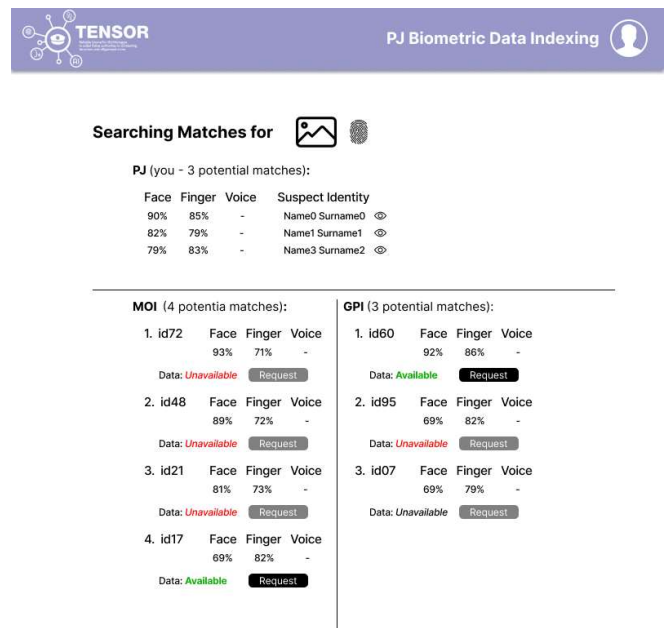


Fig. 4: Searching Results

## ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union's Horizon Europe research and innovation programme under the Grant Agreement No 101073920 (TENSOR). This publication reflects only the authors views. The European Union is not liable for any use that may be made of the information contained therein.



## REFERENCES

- [1] S. Hufnagel, C. Harfield, and S. Bronitt, *Cross-border law enforcement*. Taylor & Francis., 2012.
- [2] J. Daskal, “Law enforcement access to data across borders: The evolving security and rights issues,” *J. Nat’l Sec. L. & Pol’y*, vol. 8, p. 473, 2015.
- [3] L. Chi and X. Zhu, “Hashing techniques: A survey and taxonomy,” *ACM Computing Surveys (Csur)*, vol. 50, no. 1, pp. 1–36, 2017.
- [4] X. Luo, H. Wang, D. Wu, C. Chen, M. Deng, J. Huang, and X.-S. Hua, “A survey on deep hashing methods,” *ACM Transactions on Knowledge Discovery from Data*, vol. 17, no. 1, pp. 1–50, 2023.
- [5] K. Miniadou, A. Leonidis, G. T. Papadopoulos, and C. Stephanidis, “Enhancing secure cross-border collaboration among law enforcement agencies for facial biometric search,” in *2024 5th International Conference in Electronic Engineering, Information Technology & Education (EEITE)*. IEEE, 2024, pp. 1–6.
- [6] A. Acar, H. Aksu, A. S. Uluagac, and M. Conti, “A survey on homomorphic encryption schemes: Theory and implementation,” *ACM Computing Surveys (Csur)*, vol. 51, no. 4, pp. 1–35, 2018.
- [7] V. Rijmen and J. Daemen, “Advanced encryption standard,” *Proceedings of federal information processing standards publications, national institute of standards and technology*, vol. 19, p. 22, 2001.
- [8] R. Gutiérrez, V. Rampérez, H. Paggi, J. A. Lara, and J. Soriano, “On the use of information fusion techniques to improve information quality: Taxonomy, opportunities and challenges,” *Information Fusion*, vol. 78, pp. 102–137, 2022.
- [9] C. Yan, G. Pang, X. Bai, C. Shen, J. Zhou, and E. Hancock, “Deep hashing by discriminating hard examples,” in *Proceedings of the 27th ACM international conference on multimedia*, 2019, pp. 1535–1542.
- [10] Z. Cao, Z. Sun, M. Long, J. Wang, and P. S. Yu, “Deep priority hashing,” in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 1653–1661.
- [11] A. Tarannum, Z. U. Rahman, L. K. Rao, T. Srinivasulu, and A. Lay-Ekuakille, “An efficient multi-modal biometric sensing and authentication framework for distributed applications,” *IEEE Sensors Journal*, vol. 20, no. 24, pp. 15 014–15 025, 2020.
- [12] S. Aleem, P. Yang, S. Masood, P. Li, and B. Sheng, “An accurate multi-modal biometric identification system for person identification via fusion of face and finger print,” *World Wide Web*, vol. 23, no. 2, pp. 1299–1317, 2020.
- [13] E. mehdi Cherrat, R. Alaoui, and H. Bouzahir, “Convolutional neural networks approach for multimodal biometric identification system using the fusion of fingerprint, finger-vein and face images,” *PeerJ Computer Science*, vol. 6, p. e248, 2020.
- [14] K. N. Win, K. Li, J. Chen, P. F. Viger, and K. Li, “Fingerprint classification and identification algorithms for criminal investigation: A survey,” *Future Generation Computer Systems*, vol. 110, pp. 758–771, 2020.
- [15] D. Dessimoz and C. Champod, “A dedicated framework for weak biometrics in forensic science for investigation and intelligence purposes: The case of facial information,” *Security Journal*, vol. 29, pp. 603–617, 2016.
- [16] F. Jansen, J. Sánchez-Monedero, and L. Dencik, “Biometric identity systems in law enforcement and the politics of (voice) recognition: The case of siip,” *Big Data & Society*, vol. 8, no. 2, p. 20539517211063604, 2021.
- [17] H.-W. Ng and S. Winkler, “A data-driven approach to cleaning large face datasets,” in *2014 IEEE international conference on image processing (ICIP)*. IEEE, 2014, pp. 343–347.
- [18] L. Wolf, T. Hassner, and I. Maoz, “Face recognition in unconstrained videos with matched background similarity,” in *CVPR 2011*. IEEE, 2011, pp. 529–534.
- [19] J. Yamagishi, C. Veaux, and K. MacDonald, *CSTR VCTK Corpus: English Multi-speaker Corpus for CSTR Voice Cloning Toolkit (version 0.92)*, 2019.
- [20] D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar *et al.*, *Handbook of fingerprint recognition*. Springer, 2009, vol. 2.
- [21] M. Zhang, X. Zhe, and H. Yan, “Orthonormal product quantization network for scalable face image retrieval,” *Pattern Recognition*, vol. 141, p. 109671, 2023.
- [22] H. Zhu, M. Long, J. Wang, and Y. Cao, “Deep hashing network for efficient similarity retrieval,” in *Proceedings of the AAAI conference on Artificial Intelligence*, vol. 30, no. 1, 2016.