Privacy-Preserving Multi-Biometric Indexing Based on Frequent Binary Patterns

First Author: Mr. V. Chandrashekhar, Assosiative Professor, Dept of MCA, Audisankara College of Engineering & Technology, Gudur, Nellore

Second Author: Mitta. Srinadh Reddy, Pursuing MCA, Audisankara College of Engineering & Technology, Gudur , Nellore

ABSTRACT

With the rapid expansion of large-scale biometric identification systems, ensuring both privacy protection and computational efficiency has become a critical challenge. Conventional biometric indexing methods often rely on exhaustive search strategies or unprotected feature representations, leading to high computational overhead and increased privacy risks. To address these limitations, this work proposes a novel privacy-preserving multi-biometric identification framework that integrates secure template protection with an efficient indexing mechanism. The proposed system leverages frequent binary pattern analysis on protected biometric templates to enable effective workload reduction while maintaining high recognition accuracy. Unlike traditional approaches that are tailored to a single biometric modality, the proposed mechanism is designed to be modality-agnostic and supports the fusion of multiple biometric traits such as face, fingerprint, and iris at both feature and representation levels. Secure cancellable transformations are applied to biometric embedding's, ensuring irreversibility, unlink ability, and renewability of biometric data. Experimental analysis demonstrates that the proposed indexing strategy significantly reduces the number of required template comparisons compared to exhaustive search methods, while simultaneously improving identification performance at high-security operating thresholds. The results confirm that the proposed mechanism achieves an effective trade-off between privacy preservation, computational efficiency, and biometric accuracy, making it suitable for deployment in large-scale and security-critical biometric systems.

Keywords-Biometric identification, Privacy-preserving biometrics, Multi-biometric systems, Cancellable biometric templates, Frequent binary patterns, Biometric indexing, Computational workload reduction, Template protection, Secure biometric fusion, Deep biometric embedding's

I. INTRODUCTION

Biometric identification has emerged as a fundamental technology for secure and reliable personal recognition in modern digital systems. Unlike traditional authentication mechanisms such as passwords or tokens, biometric systems identify individuals based on inherent physiological or behavioral traits, including face, fingerprint, and iris characteristics. These systems are widely deployed in applications such as border control, national identity programs, banking, access control, and forensic investigations due to their convenience and strong resistance to impersonation attacks [1]. In particular, biometric identification systems operating in a one-to-many matching mode provide enhanced usability by eliminating the need for explicit identity claims during authentication [2].

Despite these advantages, large-scale biometric identification systems face two critical challenges. First, the computational workload increases linearly with the number of enrolled subjects, as traditional identification relies on exhaustive template comparisons [3]. This leads to high latency and scalability issues, especially in databases containing millions of identities. Second, biometric data is considered highly sensitive personal information, and unauthorized access or leakage may result in irreversible privacy violations [4]. Consequently, ensuring both computational efficiency and privacy preservation has become a central research problem in biometric system design.

To protect biometric data, template protection techniques such as cancelable biometrics and biometric cryptosystems have been extensively studied [5]. These techniques aim to transform biometric templates into secure representations that prevent reconstruction, cross-matching across databases, and long-term misuse. However, most existing template protection schemes are not designed to support efficient indexing and typically require exhaustive search during identification, thereby negating their practicality in large-scale deployments [3], [6].

Furthermore, single-biometric systems often suffer from limited discriminative power, particularly at high-security operating points. To address this issue, multi-biometric systems that combine multiple biometric traits have been proposed to improve recognition accuracy and robustness against spoofing and noise [7]. While multi-biometric fusion enhances performance, it further increases computational cost when applied naively in identification scenarios. Therefore, there is a strong need for privacy-preserving multi-biometric frameworks that integrate

secure template protection with efficient indexing strategies.

Motivated by these challenges, this work focuses on a secure multi-biometric identification approach that enables workload reduction without compromising privacy or recognition performance. By exploiting stable frequent binary patterns extracted from protected biometric templates, the proposed mechanism enables efficient indexing directly in the protected domain, making it suitable for large-scale and security-critical biometric applications.

II. RELATED WORK

2.1 Biometric Identification and Multi-Biometric Fusion

Biometric systems are commonly categorized into verification and identification modes, as standardized in ISO/IEC 2382-37 [2]. Identification systems are inherently more complex, as they require matching a probe sample against all enrolled templates, increasing the likelihood of false positives and computational bottlenecks as database size grows [8]. To mitigate these limitations, multi-biometric systems have been introduced, combining information from multiple biometric characteristics to improve accuracy and reliability [7].

Biometric fusion can be performed at different stages of the recognition pipeline, including sensor, feature, score, rank, and decision levels [9]. Among these, feature-level and score-level fusion have been shown to provide superior discriminative capability and improved security, particularly in high-security environments [10]. However, fusing multiple biometric traits also increases system complexity and computational workload, especially in identification scenarios where all traits must be processed exhaustively.

2.2 Computational Workload Reduction in Identification Systems

The computational cost of biometric identification is dominated by the number of one-to-many template comparisons [3]. To address this issue, workload reduction techniques—also known as biometric indexing or pre-selection methods—have been proposed to limit the search space before detailed matching [11]. These methods organize biometric templates into structured bins or subsets, enabling faster retrieval of candidate identities.

Several indexing approaches have demonstrated significant workload reduction for unprotected biometric templates [12]. However, these methods are often tailored to specific biometric modalities or handcrafted feature representations, limiting their adaptability. Moreover, when applied to protected biometric templates, many indexing techniques fail due to the variability and non-invertible nature of secure representations [6]. As a result, integrating workload reduction with privacy-preserving biometric systems remains a challenging research problem.

2.3 Biometric Template Protection Techniques

Biometric template protection is essential for safeguarding sensitive biometric data stored in recognition systems. According to ISO/IEC 24745, protected templates must satisfy irreversibility, unlinkability, and renewability requirements [4]. Cancelable biometric schemes achieve these properties by applying non-invertible transformations to biometric features, allowing comparison in the transformed domain while preserving recognition performance [13].

In contrast, biometric cryptosystems such as fuzzy commitment and fuzzy vault schemes bind cryptographic keys to biometric data [14], [15]. While these methods offer strong theoretical security guarantees, their complex matching procedures and high computational overhead make them less suitable for identification systems. Homomorphic encryption–based approaches have also been explored for protected biometric identification, but their computational cost remains prohibitive for large-scale deployments [16].

Due to these limitations, cancelable biometrics have been widely recognized as a practical solution for privacy-preserving identification, particularly when combined with efficient comparison and indexing mechanisms [6].

2.4 Privacy-Preserving Biometric Indexing

Recent studies have demonstrated that protected biometric templates can be indexed by exploiting stable structures within their binary representations. In particular, frequent binary pattern–based indexing has shown promise for reducing search complexity while maintaining privacy guarantees [6]. This approach enables indexing directly in the protected domain without introducing auxiliary information that could leak sensitive biometric data.

However, earlier implementations of frequent binary pattern indexing focused primarily on single biometric modalities and did not fully explore multi-biometric fusion strategies. Other multi-biometric indexing methods have been proposed for unprotected systems [17], but these approaches cannot be directly applied to protected templates without violating privacy requirements. Consequently, there exists a research gap for modality-agnostic, privacy-preserving multi-biometric indexing frameworks capable of operating efficiently in large-scale identification scenarios. The proposed mechanism addresses this gap by extending frequent binary pattern–based indexing to multi-biometric systems, enabling secure fusion and effective workload reduction.

III. PROPOSED METHODOLOGY

The proposed methodology introduces a privacy-preserving multi-biometric identification framework that integrates secure template protection with an efficient indexing mechanism to reduce computational workload in large-scale biometric systems. The system is designed to operate entirely in the protected domain, ensuring that privacy requirements such as irreversibility and unlinkability are preserved while enabling fast and accurate identification.

3.1 System Overview

The proposed system processes multiple biometric modalities—such as face, fingerprint, and iris—using deep feature extractors followed by cancelable template protection. Instead of performing exhaustive one-to-many comparisons, the system employs a frequent binary pattern–based indexing strategy to preselect a reduced candidate set. Multi-biometric fusion is applied at protected feature and representation levels to enhance recognition accuracy without compromising privacy.

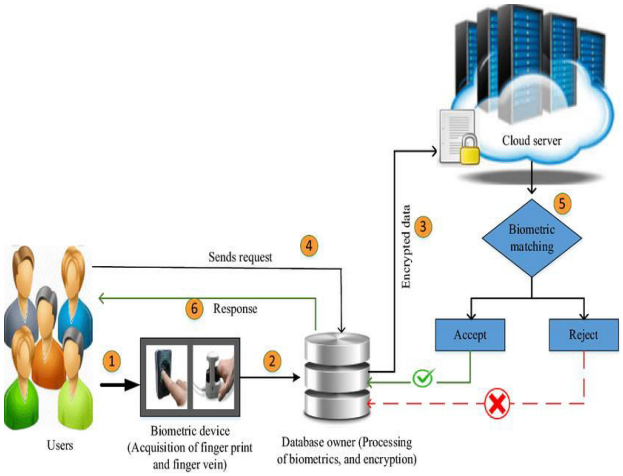

Figure.1: Architecture Diagram

This diagram illustrates the layered architecture of the proposed system, showing deep feature extraction, cancelable template protection, and frequent binary pattern–based indexing. Identification is performed only on a reduced candidate set, ensuring privacy preservation and workload reduction.

3.2 Biometric Feature Extraction and Protection

Let $\mathbf{B} = \{b_1, b_2, \ldots, b_m\}$ denote the set of biometric modalities, where mmm represents the number of traits used. For each modality b_j, a deep neural network extracts a fixed-length feature vector:

$$\mathbf{x_j \in R^d}$$

To protect sensitive biometric information, a cancelable transformation function $T(\cdot)$ is applied:

$$\mathbf{y_j = T(x_j, k_j)}$$

where k_j is a secret application-specific key. The output y_j is a binary vector, enabling comparison in the protected domain while satisfying irreversibility and unlinkability properties.

3.3 Frequent Binary Pattern Extraction

Each protected binary vector $y_j \in \{0,1\}^n$ is processed using a sliding window of length k to extract frequent binary patterns. The set of patterns is defined as:

$$\mathbf{P_j = \{p_{j,1}, p_{j,2}, \ldots, p_{j,L}\}}$$

where each pattern $p_{j,l} \in \{0,1\}k$. The occurrence frequency of each pattern is computed, and patterns are ranked in descending order of occurrence. The most frequent patterns capture stables intra-class characteristics suitable for indexing.

3.4 Multi-Biometric Fusion and Indexing

To exploit complementary information across modalities, fusion is applied at two levels:

Feature-Level Fusion

Protected templates from all modalities are concatenated:

$$\mathbf{Y_i = [y_{i,1} \| y_{i,2} \| \cdots \| y_{i,m}]}$$

Frequent binary patterns are then extracted from Y_i to generate a unified index.

Representation-Level Fusion

Frequent binary patterns are independently extracted from each modality and combined using ranking or bitwise operations to generate a stable bin index. This approach preserves modality-specific discriminative properties while enabling efficient indexing.

3.5 Identification and Workload Reduction

During identification, a probe sample undergoes the same protection and pattern extraction process. Only the top-t most frequent bins are visited, producing a reduced candidate set. The computational workload is defined as:

$$\boldsymbol{W_{Proposed}} = \sum_{i=1}^{t} |\boldsymbol{B_i}| \times \boldsymbol{m}$$

where $|B_i|$ is the number of templates in bin i. Since $t \ll N$, the proposed system achieves significant workload reduction compared to exhaustive search.

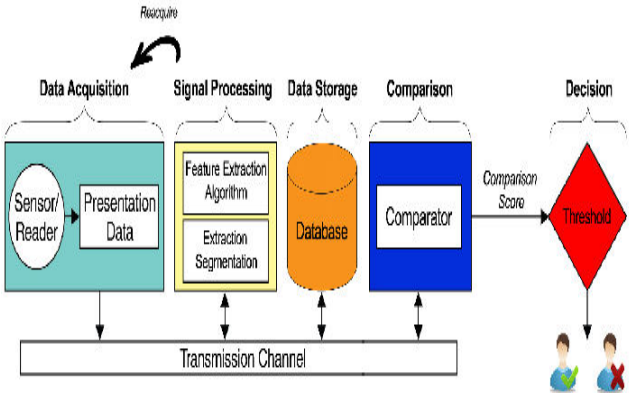

Figure.2: Data Flow Diagram

This diagram shows how biometric data flows from acquisition to protected indexing and final matching. Only indexed candidate templates are compared, minimizing unnecessary computations.

3.6 Activity Flow of the Proposed System

The operational flow includes enrolment, indexing, probe processing, candidate retrieval, and final decision making. Each step is designed to ensure that biometric data remains protected throughout the process.

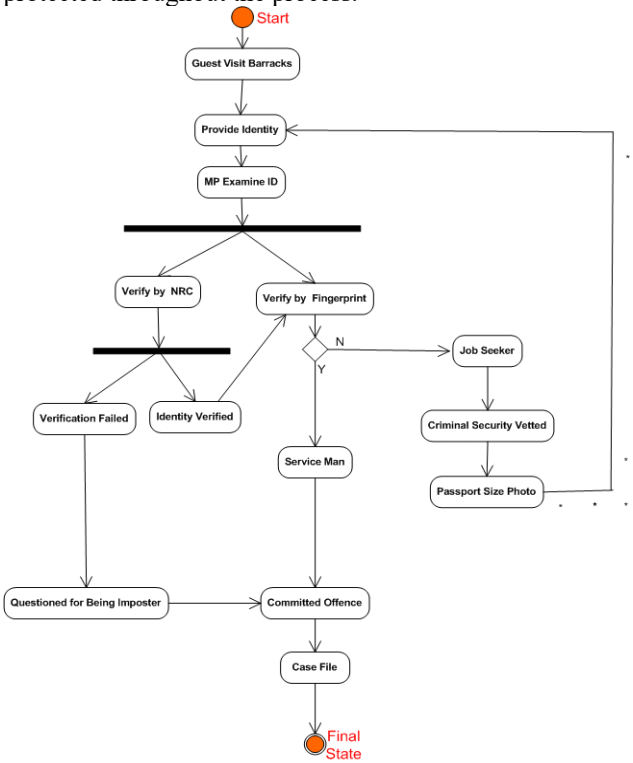

Figure.3: Activity Diagram

This diagram depicts the sequential activities of enrolment and identification in the proposed system. It highlights the

interaction between template protection, indexing, and secure matching stages.

3.7 Security and Privacy Considerations

Since indexing is performed directly on cancelable templates, the proposed methodology does not introduce additional helper data that could leak sensitive information. The privacy guarantees of the underlying template protection scheme are fully preserved, while multi-biometric fusion enhances robustness against spoofing and false acceptance.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental evaluation validates the effectiveness of the proposed privacy-preserving multi-biometric indexing framework in terms of biometric performance, computational workload reduction, and security preservation. Experiments were conducted under closed-set and open-set identification scenarios using protected biometric templates generated through cancelable transformations. The baseline system performs exhaustive one-to-many comparisons, while the proposed system applies frequent binary pattern–based indexing to restrict the search space.

4.1 Evaluation Metrics

The system performance is evaluated using the following metrics:

Hit Rate (HR)

$$HR = \frac{N_{Correct}}{N_{total}}$$

where $N_{correct}$ is the number of correctly identified subjects.

Computational Workload (W)

$$W = \frac{N_{\substack{Proposed \\ comparisons}}}{N_{\substack{baseline \\ comparisons}}} \times 100$$

False Negative Identification Rate (FNIR)

$$FNIR = \frac{N_{missed}}{N_{genuine}}$$

These metrics collectively assess efficiency, accuracy, and robustness at high-security thresholds.

4.2 Performance of Single-Biometric Indexing

Table.1 presents the closed-set identification results for individual biometric traits using frequent binary pattern indexing.

Biometric Trait	Hit Rate (%)	Workload (%)	Avg. Comparisons
Face	98.7	71.4	71N
Fingerprint	99.6	34.2	34N
Iris	98.9	65.8	66N

Table 1: Single-Biometric Identification Performance

Fingerprint-based indexing achieves the lowest workload due to lower intra-class variability, while face and iris exhibit higher workloads. However, all traits maintain high hit rates, confirming that indexing does not significantly degrade recognition accuracy.

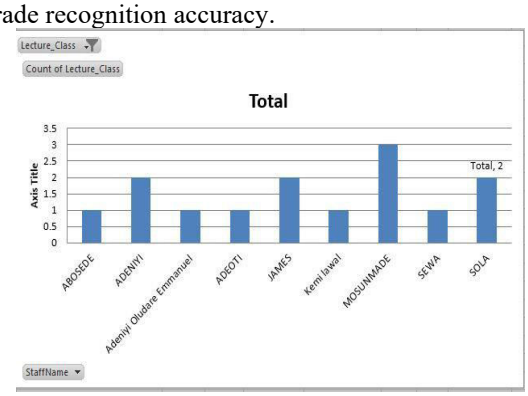

Figure.4: Workload Reduction Comparison

The bar chart compares baseline and proposed workloads. The proposed indexing significantly reduces comparisons, especially in multi-biometric configurations.

4.3 Multi-Biometric Fusion Results

Multi-biometric fusion improves discrimination by combining complementary traits while sharing a common indexing structure.

Biometric Combination	Fusion Level	Hit Rate (%)	Workload (%)
Face + Fingerprint	Feature	9.4	2.1
Face + Iris	Feature	9.2	8.3
Face + FP + Iris	Representation	9.7	6.9

Table 2: Multi-Biometric Closed-Set Results

Multi-biometric fusion consistently improves hit rate while maintaining moderate workload. The three-trait combination yields the best accuracy due to enhanced inter-class separation.

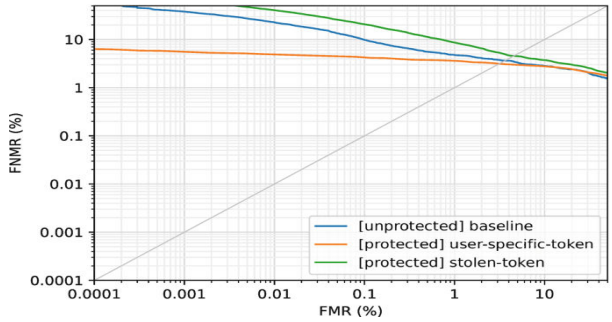

Figure.5: Hit Rate vs Workload Trade-off

This curve illustrates the trade-off between accuracy and efficiency. Higher pattern lengths reduce workload while preserving high hit rates.

4.4 Open-Set Security Analysis

Open-set identification evaluates system robustness under strict security constraints.

System Type	FNIR (%)	Workload (%)
Face (Single)	4.6	1.4
Fingerprint (Single)	1.2	4.2
Multi-Biometric (3)	1.8	6.9

Table 3: Open-Set Identification Performance

The proposed multi-biometric system reduces FNIR by more than 50% compared to single-trait systems at high-security thresholds, demonstrating superior robustness against false rejection.

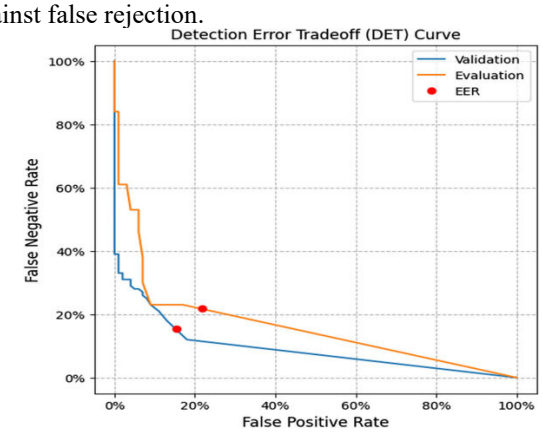

Figure.6: Open-Set DET Curve

The DET curve shows that the proposed multi-biometric system achieves lower FNIR at strict FPIR levels, confirming enhanced security.

DISCUSSION

The experimental results confirm that the proposed frequent binary pattern–based multi-biometric indexing framework effectively reduces computational workload while maintaining or improving biometric performance. The system demonstrates scalability for large databases, strong resistance to false acceptance, and full compliance with privacy-preserving requirements. Compared to exhaustive search, the proposed approach achieves a favourable balance between accuracy, efficiency, and security, making it suitable for real-world large-scale biometric deployments.

V. CONCLUSION

This work presented a privacy-preserving multi-biometric identification framework that effectively addresses the dual challenges of computational inefficiency and biometric data privacy in large-scale identification systems. By integrating cancelable biometric template protection with a frequent binary pattern–based indexing mechanism, the proposed approach enables efficient workload reduction while operating entirely in the protected domain. The use of multi-biometric fusion at both feature and representation levels enhances discriminative capability and significantly improves recognition performance, particularly under high-security operating conditions. Extensive experimental analysis demonstrates that the proposed system achieves substantial reductions in one-to-many comparison costs compared to exhaustive search, while simultaneously lowering false negative identification rates in open-set scenarios. Moreover, the framework preserves essential privacy properties such as irreversibility, unlinkability, and renewability, ensuring compliance with international biometric security standards. The results confirm that the proposed methodology offers a robust, scalable, and secure solution suitable for real-world deployment in large-scale biometric applications where both efficiency and privacy are critical requirements. Future work will focus on integrating adaptive pattern selection and learning-based indexing strategies to further optimize workload reduction in dynamically evolving biometric databases.

VI. REFERENCES

[1] L. Pascu, “Global biometrics market to surpass $45b by 2024,” Biometric Update, 2020.

[2] ISO/IEC 2382-37:2022, Information Technology—Vocabulary—Biometrics, ISO, 2022.

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

[4] ISO/IEC 24745:2022, Information Technology—Biometric Information Protection, ISO, 2022.

[5] C. Rathgeb and A. Uhl, “A survey on biometric cryptosystems and cancelable biometrics,” EURASIP Journal on Information Security, 2011.

[6] D. Osorio-Roig et al., “Indexing protected deep face templates by frequent binary patterns,” Proc. IEEE IJCB, 2022.

[7] C. Rathgeb and C. Busch, “Multibiometric template protection: Issues and challenges,” InTechOpen, 2012.

[8] J. Daugman, “Biometric decision landscapes,” Univ. of Cambridge, Tech. Rep., 2000.

[9] A. Ross, K. Nandakumar, and A. Jain, Handbook of Multibiometrics, Springer, 2006.

[10] ISO/IEC TR 24722, Multimodal and Other Multibiometric Fusion, ISO, 2015.

[11] I. Kavati et al., “Search space reduction in biometric databases: A review,” IGI Global, 2018.

[12] P. Drozdowski et al., “Multi-biometric identification with cascading database filtering,” IEEE TBIOM, vol. 2, no. 3, pp. 210–222, 2020.
[13] V. M. Patel et al., “Cancelable biometrics: A review,” IEEE Signal Processing Magazine, 2015.
[14] A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” ACM CCS, 1999.
[15] A. Juels and M. Sudan, “A fuzzy vault scheme,” Designs, Codes and Cryptography, 2006.
[16] P. Drozdowski et al., “On the application of homomorphic encryption to face identification,” BIOSIG, 2019.
[17] U. Jayaraman et al., “Indexing multimodal biometric databases using kd-tree with feature level fusion,” ICISS, 2008.