Indian Journal of Signal Processing (IJSP)
ISSN: 2582-8320 (Online), Volume-5 Issue-4, November 2025

Adaptive Multimodal Biometric Recognition Framework Integrating Iris and Fingerprint Modalities for Robust and Interpretable Authentication

S. Ramesh, V. Krishnaveni

Abstract: ***Biometric recognition is a vital technology for secure identification. However, unimodal systems often face several drawbacks, including reduced reliability in challenging demographic groups such as children, environmental constraints, and susceptibility to spoofing. To address these problems, this work presents a scalable multimodal biometric architecture that integrates fingerprint and iris modalities to enhance identification accuracy, resilience, and interpretability. In the proposed architecture, unique iris patterns are captured using a Multi-Layer Perceptron Mixer (MLP-Mixer), and concise yet discriminative fingerprint attachments are generated using a Variational Autoencoder (VAE). Using a Triplet Network to improve the distinction between real and fake samples, the matching performance is further strengthened. Appropriate blending is achievd through feacture-level fusion using a Cross-Attention Transformer, which dynamically aligns complementary iris and fingerprint embeddings. Crucially, by highlighting each feature's contribution to the final decision, Integrated Gradient (IG) are used to guarantee integritu and openness.. The efficacy of this technique is demonstrated by experiments using benchmark iris and fingerprint datasets, which achieved an overall identification accuracy of 97.8%, with Equal Error Rates (EER) of 0.6% for iris and 0.7% for fingerprints. The robustness of the suggested paradigm is demonstrated by comparison against unimodal and multimodal baselines, especially in situations involving noisy data and early-age biometric identification. These results highlight the system's usefulness for applications in safeguarding children, scalable authentication, and encrypted control of entry. In the final analysis, this study not only creates a new path for interpretable multisensory fusion but also creates the foundation for extending biometric solutions to larger populations and operational settings.***

Keywords: ***CASIA Iris Datasets, NIST Special Databases, MLP-Mixer, Variational Autoencoders, Triplet Network, Cross-Attention Transformers, Integrated Gradients.***

Nomenclature:
IG: Integrated Gradients
VAE: Variational Autoencoder
CASIA: Chinese Academy of Sciences Institute of Automation
NIST: National Institute of Standards and Technology
MLP Mixer: Multi-Layer Perceptron Mixer
EER: Equal Error Rate
FAR: False Acceptance Rate
FRR: False Rejection Rate
RMSE: Root Mean Square Error
IG: Integrated Gradients
DNSN: Deep Noise Suppression Network
NIR: Near-Infrared
TPR: True Positive Rate
TNR: True Negative Rate

Manuscript received on 23 September 2025 | First Revised Manuscript received on 05 October 2025 | Second Revised Manuscript received on 18 October 2025 | Manuscript Accepted on 15 November 2025 | Manuscript published on 30 November 2025.

* Correspondence Author(s)
S. Ramesh*, Research Scholar, Department of Electronics and Communication Engineering, PSG College of Technology, Coimbatore (Tamil Nadu), India. Email ID: rameshsivagaminathan@gmail.com, ORCID ID: 0009-0000-0041-1860
Dr. V. Krishnaveni, Professor & Head, Department of Electronics and Communication Engineering, PSG College of Technology, Coimbatore (Tamil Nadu), India. Email ID: vk.ece@psgtech.ac.in, ORCID ID: 0000-0002-4526-0055

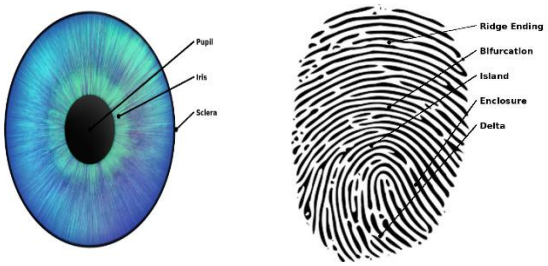
I. INTRODUCTION

Establishing reliable and precise identity verification has become crucial in today's increasingly secure digital environment. As a result, biometric recognition has become an essential technology, and increasingly multimodal approaches—especially those that incorporate fingerprint and iris characteristics—are gaining popularity. Multimodal frameworks leverage the potentially beneficial effects of various physiological indicators, in contrast to unimodal systems that are often limited by noise sensitivity or susceptibility to spoofing. This integration enhances endurance in challenging circumstances, improves identification accuracy, and strengthens defences against fraudulent efforts. The following basic criteria have been frequently employed to construct biometric identifiers: collectability (convenient ease of acquisition), permanence (durability over time), distinctiveness (specificity to each individual), and universality (accessibility across all persons). Both iris and fingerprint are the most notable on this band due to their exceptional degrees of temporal stability and uniqueness, which make them perfect for high-assurance identification systems.

Early unimodal models, which relied solely on characteristics such as speech or fingerprints, have given way to sophisticated multimodal systems that can function effectively in a variety of settings and demographics. Advances in pattern recognition techniques, predictive algorithms, and sensors have all contributed to this change. To enable sustainable implementation in large-scale, practical applications, contemporary biometric networks have been developed to surpass existing identification efficiency and accommodate demographic variations.

The iris and fingerprint have the most discriminative power among physiological identifiers. The iris, a

collection of closely related individuals, is a pigmented, contractile tissue that surrounds the pupil. It features intricate structures such as radial furrows, crypts, and concentric patterns that are incredibly unique. Its resilience to external factors and ageing further supports its value for long-lasting recognition. Similar to this, there are essentially three main types of fingerprints that can be recognised during pregnancy and identified by their ridge structures: loops, whorls, and arches. Arches appear as straightforward wave-like crests without looping, whorls take the form of spiral-like circular ridges, and loops recur onto the radial or upper edge. Fingerprints have long been a mainstay of both forensic and civilian identification verification due to their anatomical persistence.



[Fig.1: Anatomical Regions of the Human Eye and Fingerprint Patterns]

The three conventional fingerprint configurations—loops, whorls, and arches—are shown in Fig. 1(b). In contrast, Fig. 1(a) presents an integrated view illustrating the structural areas of the human eye, focusing on the iris. This example illustrates the inherent constancy and distinctiveness of these traits, which form the biological basis for the proposed multimodal biometric recognition system.

The proposed study presents a strong and flexible multimodal framework intended to provide scalable identity identification in a variety of operational environments by utilising the complementary capabilities of fingerprint and iris characteristics. The technology ensures that decisions are interpretable, particularly addressing issues such as age-induced variances and learning discrepancies. As seen, our study not only fortifies the cornerstones of safe biometric identification but also paves the way for sophisticated, open, and deployable identity verification systems in both high-security and civilian settings.

II. LITERATURE SURVEY

With a particular focus on resolving age-related variability and developmental changes in physiological features, biometric recognition research has made significant strides. A thorough analysis of fusion methods in multimodal biometric devices was provided by Bala et al. [1], who placed special emphasis on algorithmic approaches and performance trade-offs. In their discussion of developments in multimodal biometric authentication, Pahuja and Goel [2] emphasised system comparisons and new security concerns. A thorough analysis of fingerprint biometric authentication systems was conducted by Sumalatha et al. [3], who highlighted weaknesses, template protection, and fusion strategies. Ahmed et al. [4] provided an overview of combined iris and fingerprint systems, demonstrating complementary strengths and integration frameworks. Al-Dabbas et al. [5] conducted a brief survey on multimodal systems that utilise face and fingerprint recognition, reporting empirical findings on performance benchmarks.

Ryu et al. [6] analysed continuous multimodal authentication methods, emphasising the need for stability and real-time verification. Almuwayziri et al. [7] explored fingerprint–vein biometric fusion with deep learning, combining survey analysis with empirical evaluation. Nguyen et al. [8] provided a broad review of deep learning approaches for iris recognition, identifying challenges in generalisation and cross-domain adaptation. Shaheed et al. [9] surveyed physiological biometrics, situating fingerprint and iris studies within a broader context of traits such as ECG and gait. Khade et al. [10] reviewed iris liveness detection techniques, focusing on anti-spoofing countermeasures and future research directions.

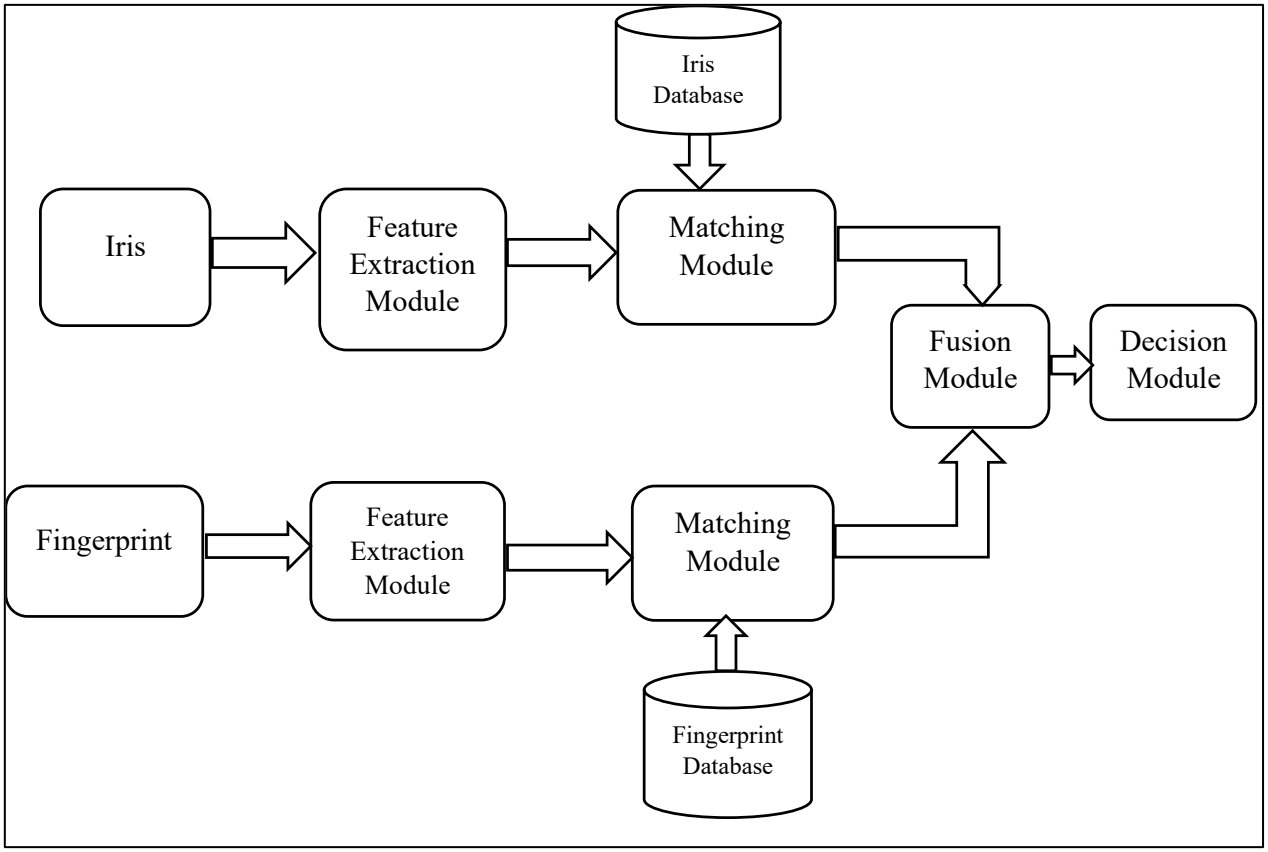
Deepika et al. [11] examined multimodal biometrics involving palm and wrist traits, with emphasis on optimised feature extraction strategies. Singh et al. [12] surveyed multimodal presentation attack detection, identifying material-based spoofing challenges and highlighting multimodal fusion as a mitigation strategy. Meiramkhanov et al. [13] provided a comprehensive exploration of biometric authentication techniques, with an emphasis on security frameworks in cloud-centric systems. Convolutional neural networks have been shown to enhance multimodal person identification through hybrid fusion, as demonstrated by Jayanna and Yadava [14]. By examining the weaknesses presented in different spoofing materials, Agarwal [15] investigated fusion in multimodal spoofing attacks.

A comprehensive review of iris segmentation and identification techniques was presented by Rasheed et al. [16], who emphasised the contribution of deep learning to enhancing the robustness of segmentation. A thorough analysis of biometric systems based on psychological and biological characteristics was offered by Liu et al. [17], who also highlighted trade-offs between security and privacy. In their analysis of biometric security options and challenges, Arora and Bhatia [18] addressed template protection and the difficulties associated with large-scale adoption. In their study of secure multimodal systems, Akulwar and Vijapur [19] offered methods for incorporating cryptography into multimodal structures. Lastly, a deep learning-based face-iris multimodal detection system was demonstrated by Hattab and Behloul [20], who reported significant improvements in precision for both classification and feature learning. In addition to highlighting the broader history of biometric progress across various paradigms and eras, these investigations collectively have strengthened the iris as the standard in biometric uniqueness and permanence.

III. PROPOSED METHODOLOGY

Overall, the innovative architecture for bidirectional biometric identification presented in this study combines the discriminative advantages of fingerprints and retinal modes to provide remarkable accuracy and resilience. The systematic approach, which utilises sophisticated algorithms at all stages of the recognition process, is

illustrated in Fig. 2. It aims to achieve consistent results across various ethnic groupings.



[Fig.2: Systematic Workflow for Iris and Fingerprint-Based Multimodal Biometrics]

A next-generation multimodal biometric framework, which integrates fingerprint and iris features into a unified recognition system, is presented in the process illustration. Acquisition is the initial step, where carefully selected data guarantees accuracy and variety in the inputs that are recorded. The fingerprint modality is encoded using VAEs to provide compact, noise-resilient representations of ridge and minutiae patterns. In contrast, the iris modality is managed using an MLP-Mixer to capture both global and local textural characteristics during feature extraction. A Triplet Network promotes accuracy by comparing anchor, positive, or adverse data via the Pairing Section, which further refines these embeddings. A Cross-Attention Transformer, which dynamically aligns inter-modality relationships and maximises the contribution of each biometric source, is then used to merge the multimodal embeddings in the Fusion Module. Subsequently, the Decision Module ensures accountability and confidence in network predictions by employing IG, which provides interpretability by assigning prediction results to modality-specific attributes. When taken as a whole, this approach offers a reliable, flexible, and understandable method of biometric recognition.

A. Acquisition Stage: Leveraging Biometric Repositories for Multimodal Input

The acquisition module combines two well-known repositories, the CASIA Iris Dataset and the NIST Special Fingerprint Databases, to guarantee high-quality input at the foundational stage. By combining a variety of subject information and acquiring scenarios, these datasets collectively provide a thorough, universally applicable foundation for multimodal technologies.

Approximately 1,000 photos taken under various lighting and demographic conditions comprise the CASIA Iris Dataset, a standard in iris identification research. While the high-resolution picture captures complex iris textures, which are necessary for accurate feature extraction, this variety allows for efficient modelling of intra- and inter-subject variability. On the other hand, NIST Special Databases, which provide grayscale scans of fingerprints at 300 dpi resolution, constitute the standard for fingerprinting research. These datasets provide an accurate basis for constructing robust identification algorithms, as they encompass age-related impairments, pressure-induced distortions, and natural ridge changes.

The framework creates an integrated input zone by combining these archives and harmonising cognitive (fingerprint) and physiological (iris) identities. Consistency and independence in detection are improved by incorporating it across demographic variations and environmental variances. This integration enhances recognition reliability and adaptability across both environmental variability and demographic diversity. Consistency and flexibility in identification are enhanced by incorporating it across ethnic diversity and ecological variation. Additionally, the authenticity and diversity of those datasets fortify the later preprocessing and feature extraction phases, establishing the foundation for a scalable multimodal recognition system.

Table I: Dataset Specifications for Multimodal Biometric Acquisition

Dataset Name	Biometric Modality	Image Resolution	Sample Size	Subjects	Acquisition Conditions	Distinctive Features
CASIA Iris v3	Iris	320 × 280 pixels (grayscale)	1000 iris images	500	Near-Infrared (NIR) lighting, indoor settings	High-quality texture patterns, controlled dilation, and noise-free images
NIST Special Database 14	Fingerprint	500 dpi (grayscale scans)	1000 fingerprint images	500	Rolled and flat prints, multi-session captures	Real-world variability, pressure distortions, and age-diverse subjects

The datasets used in this framework are compiled in Table 1, with a focus on modality, resolution, and acquisition conditions. A diverse variety with dependable output is ensured throughout the two modes when CASIA Iris v3 and NIST SD14 are used in tandem.

i. Mathematical Description of Dataset Configuration

The multimodal dataset utilised in this approach may be formalised as follows:

$$D = \{(x_i^{iris}, x_i^{fp}, y_i)\}_{i=1}^{N} \quad \ldots \quad (1)$$

where $x_i^{iris} \in R_{320\times280}$ and $x_i^{fp} \in R_{500\times280}$ denote grayscale iris and fingerprint images respectively, and $y_i \in \{1, 2, \ldots, 500\}$ represents the subject label. The dataset comprises $N = 1000$ samples per modality, ensuring a class-balanced configuration. Each biometric class is uniformly represented:

$$\sum_{i=1}^{500} n_i^{iris} = \sum_{i=1}^{500} n_i^{fp} = 1000 \quad \ldots \quad (2)$$

With $n_i^{modality} = 2$ samples per subject. During feature extraction, iris and fingerprint data are projected to a common latent embedding space using modality-specific functions:

$$\phi^{iris}: R^{320*280} \rightarrow R^d \quad \ldots \quad (3)$$

$$\phi^{fp}: R^{500*280} \rightarrow R^d \quad \ldots \quad (4)$$

(Note: The fingerprint images are standardized to 500x280 pixels, consistent with the iris samples.)

Where d denotes the dimensionality of the learned feature representation, this representation facilitates downstream matching, fusion, and decision-making processes. All biometric samples were normalised to a standardised intensity range before embedding, ensuring uniform feature scaling across modalities. The balanced dataset configuration minimises bias, thereby facilitating fair and reliable multimodal learning in subsequent preprocessing and feature extraction stages.

B. Preprocessing, Segmentation, and Normalisation for Enhanced Feature Integrity

The biometric inputs used in this framework—iris and fingerprint images—are obtained from the curated repositories described in the acquisition stage, namely the CASIA-IrisV3 database and the NIST Special Database 14. To ensure consistency, clarity, and readiness for feature extraction, the raw inputs undergo a structured preprocessing pipeline that includes noise reduction, quality evaluation, region segmentation, and normalisation.

This phase involves reducing unwanted artefacts in ocular metrics to eliminate sensor-induced distortions, specular reflections, and inconsistent lighting. To mitigate distortion while preserving the complex textural patterns required for accurate iris detection, a Deep Noise Suppression Network (DNSN) is employed. This circular iris area is then isolated using segmented algorithms that eliminate reflected light, eyelid occlusions, and eyelash occlusions. To achieve consistency to scale, rotation, and pupil dilation, a segmented eye can be extracted into a normalised circular orientation area. This conversion makes sure that iris patterns with different structural characteristics are all projected into the same analytical domain.

Fingerprint images undergo a parallel enhancement process. Initially, ridge refinement filters improve the clarity of minutiae and local ridge flow. The core region is then extracted to focus analysis on the most discriminative areas. A quality-control module discards low-contrast or distorted samples, ensuring that only valid inputs proceed through the pipeline. The accepted patches are resized to a uniform resolution and intensity-normalised. Statistical normalisation is applied using:

$$X_{norm}(i,j) = \frac{X(i,j) - \mu}{\sigma} \quad \ldots \quad (5)$$

Where μ and σ stand for the image's mean and standard deviation, respectively, and $X(i, j)$ for the pixel intensity. By reducing brightness overall, lighting disparities are normalised, thereby optimising sample comparability.

Box 1. Pseudocode: Preprocessing, Segmentation, and Normalisation

Input: Iris image I from CASIA, Fingerprint image F from NIST
Output: Normalised iris I_{norm}, Normalised fingerprint F_{norm}
Step 1: Apply noise attenuation:
$I_{clean} \leftarrow$ SuppressNoise(I)
$F_{clean} \leftarrow$ SuppressNoise(F)
Step 2: Segment regions of interest:
$R_{iris} \leftarrow$ ExtractIris (I_{clean})
$R_{fp} \leftarrow$ ExtractCore (F_{clean})
Step 3: Apply quality assessment:
If Quality(R_{iris}) < threshold → discard
If Quality(R_{fp}) < threshold → discard
Step 4: Normalise geometric structure:
$I_{norm} \leftarrow$ PolarNormalize (R_{iris})
$F_{norm} \leftarrow$ ResizeAndEnhance (R_{fp}, N × N)

This integrated preprocessing pipeline ensures that both iris and fingerprint inputs are noise-reduced, geometrically

standardised, statistically normalised, and quality-verified. By preserving structural integrity and discarding unreliable samples, this stage establishes a robust foundation for extracting high-fidelity embeddings in the subsequent feature extraction module.

C. Advanced Feature Extraction: MLP-Mixer for Iris and VAEs for Fingerprint

Following the preprocessing stage, the normalised iris and fingerprint images are passed to the feature extraction module, where identity-relevant information is transformed into compact and discriminative embeddings. This stage is pivotal, as the fidelity of extracted features directly governs the accuracy and robustness of downstream matching and fusion operations.

For iris recognition, this work employs the MLP-Mixer, a lightweight yet powerful architecture that differs from both convolutional and transformer-based designs. Each normalised iris image is divided into non-overlapping patches, flattened, and projected into an embedding space. Let $x_i \in \mathbb{R}^{N \times D}$ represent the sequence of N patches, each with embedding dimension *D*. The token-mixing MLP captures spatial dependencies across patches as:

$$U = X + W_2\sigma(W_1X^T)^T \quad \ldots \quad (6)$$

Where W_1, W_2 are learnable weight matrices and $\sigma(\cdot)$ is a non-linear activation. The channel-mixing MLP then captures correlations across feature dimensions:

$$Y = U + V_2\sigma(V_1\,U) \quad \ldots \quad (7)$$

With V_1, V_2 as learnable weights. This two-step mechanism enables the MLP-Mixer to effectively capture both fine-grained iris patterns and long-range dependencies, producing robust embeddings that preserve distinctiveness and interpretability.

For fingerprint recognition, VAEs are adopted to encode ridge and minutiae structures within a probabilistic framework. Unlike deterministic encoders, VAEs map each fingerprint into a latent distribution N (μ, σ^2), where μ and σ^2 represent the outputs of the encoder. The latent embedding is obtained via the reparameterization trick:

$$z = \mu + \sigma \odot \varepsilon,\ \varepsilon \sim N(0, I) \quad \ldots \quad (8)$$

This formulation ensures that the latent vector z not only captures discriminative ridge structures but also models intra-class variability and acquisition-induced distortions. During training, reconstructed fingerprints from the decoder enforce compact yet informative representations, while at inference time, z is used as the feature embedding.

By combining the MLP-Mixer and VAE in a dual-stream architecture, the framework extracts noise-resilient, compact, and semantically rich modality-specific features. These embeddings retain the essential identity cues of iris and fingerprint modalities, ensuring discriminative power for subsequent matching and fusion stages.

Box 2. Pseudocode: Feature Extraction from Iris and Fingerprint Modalities

Input: Normalised iris image I, normalised fingerprint image F (from preprocessing stage)
Output: Feature embeddings E_{iris} and E_{fp}
Process:
Step 1: Divide I into N non-overlapping patches: {P_1, P_2, ..., P_n}
Step 2: Flatten patches and project into embeddings: $x_i \in \mathbb{R}^D$
Step 3: Apply token-mixing MLP → $U = X + W_2\sigma(W_1X^T)^T$
Step 4: Apply channel-mixing MLP → $Y = U + V_2\sigma(V_1\,U)$
Step 5: Aggregate outputs → $E_{iris} = Y \in \mathbb{R}^D$
Step 6: Encode fingerprint image F using VAE → μ, σ = Encoder(F)
Step 7: Sample latent vector z from N (μ, σ²): z = μ + σ ⊙ ε, ε ~ N (0, I)
Step 8: Reconstruct fingerprint F' = Decoder(z) (used only during training)
Step 9: Assign latent vector as fingerprint embedding E_{fp} = z ∈ $\mathbb{R}^D$
Step 10: Return (E_{iris}, E_{fp})

D. Precision-Driven Matching: Triplet Network for Enhanced Similarity Assessment

In the matching stage, the modality-specific embeddings derived from the iris and fingerprint streams are compared to establish identity congruence. This framework employs a Triplet Network, an extension of the Siamese paradigm, designed to optimise similarity learning through comparative analysis. Instead of evaluating pairs, the network processes triplets consisting of an anchor embedding, a positive embedding (with the same identity), and a negative embedding (with a different identity). This arrangement enforces discriminability by simultaneously minimising the distance between anchor–positive pairs and maximising the separation between anchor–negative pairs.

Formally, let $E_a, E_p, E_n \in \mathbb{R}^D$ represent the embeddings of anchor, positive, and negative samples. The objective is expressed through the triplet loss function:

$$L_{triplet} = \max\left(0, \| E_a - E_p \|_2^2 - \| E_a - E_n \|_2^2 + \alpha\right) \quad \ldots \quad (9)$$

Where this factor $\alpha > 0$ ensures that both positive and negative pairings are kept at least a certain distance apart. This approach ensures that even though embeddings from various individuals are well-separated in the latent space, representations from a single identity are grouped. There are multiple benefits to using a triplet network. In addition to effectively addressing overall ethnic contextual variability, it enhances resistance to minor intra-class differences, such as slight iris crypt deformations or fingerprint ridge pressure discrepancies. Subsequently, it is also, easily scalable thanks to its framework, which facilitates integration with other biometric features in multimodal systems.

Box 3. Pseudocode: Matching Mechanism Using Triplet Network

Input: Embeddings $E_a, E_p, E_n \in \mathbb{R}^D$ (anchor, positive, negative)
Output: Match/non-match decision for anchor sample
Process:
Step 1: Pass E_a, E_p, E_n through identical subnetworks → H_a=f(E_a), H_p=f(E_p), H_n=f(E_n)
Step 2: Compute squared distances: $D_{ap} = \|H_a - H_p\|^2$, $D_{an} = \|H_a - H_n\|^2$
Step 3: Evaluate triplet loss:
$L_{triplet} = \max(0, D_{ap} - D_{an} + \alpha)$
Step 4:
If $D_{ap} \leq$ T and $D_{an} >$ T→ return 1 (Match)
Else → return 0 (non-match)

This triplet-based approach guarantees that embeddings are optimal for relational location within the field of embedding and for bilateral evaluation. Higher discrimination, scalability, and resilience are thus achieved by the system, thereby enhancing the dependability of multimodal biometric verification.

E. Fusion Module: Adaptive Multimodal Integration via Cross-Attention Transformers

The suggested framework's integrating centre, its Fusion Module, is where the modality-specific embeddings from the fingerprints, together with the iris streams, are merged and then aligned. This step utilises a Cross-Attention Transformer that interactively develops interactions between the two forms, in contrast to conventional fusion approaches such as naïve splicing or averaging. The methodology preserves identity-relevant signals by downweighting smaller or noisy locations and emphasising complementary traits by utilising focus across modalities.

The iris embedding is denoted as $E_{iris} \in R^D$ and the fingerprint embedding is represented as $E_{fp} \in R^D$. Modality-specific alignments are made possible by the fusion process, which uses one embedding as the query and the other as the source of key-value pairs. Formally, the cross-attention operation is as follows:

$$Z = Softmax\left(\frac{Q K^T}{\sqrt{d_k}}\right) V \quad \ldots \quad (10)$$

Where $Q = W_Q E_{iris}$, $K = W_K E_{fp}$, and $V = W_V E_{fp}$, with W_Q, W_K, W_V representing learnable projection matrices. This approach creates a merged model that reflects regional individuality or broad multimodal coherence by allowing the iris embedding to focus on salient fingerprint characteristics (or vice versa).

The fused vector $V_{fused} \in \mathbb{R}^K$ gets better by stacking several cross-attention layers, where K is the size of the multimodal embedding space. This adaptive integration improves robustness across varying acquisition settings by mitigating frequent problems like pressure-induced fingerprint distortion and partial obstruction in iris scans.

Box 4. Pseudocode: Multimodal Feature Fusion via Cross-Attention Transformer

Input: $E_{iris} \in \mathbb{R}^D$ (Iris Embedding), $E_{fp} \in \mathbb{R}^D$ (Fingerprint Embedding)
Output: Fused Representation $V_{fused} \in \mathbb{R}^K$
Process:
Step 1: Project embeddings into query, key, value spaces:
$Q = W_Q E_{iris}$, $K = W_K E_{fp}$, $V = W_V E_{fp}$
Step 2: Compute cross-attention:
$Z = Softmax\left(\frac{Q K^T}{\sqrt{d_k}}\right) V$
Step 3: Apply residual connection and feed-forward MLP:
$Z' = \text{MLP}(Z + E_{iris})$
Step 4: Repeat for multiple stacked layers to refine representation
Step 5: Aggregate outputs → fused embedding $V_{fused} \in R^K$
Step 6: Pass V_{fused} to the Decision Module for final verification

By adjusting distorted and inadequate samples, this attention-driven merging technique ensures that both paradigms provide the best possible contribution to the decision-making process. Instead of depending on static rules, the algorithm develops adaptive positioning that improves recognition accuracy, interpretability, and resilience under a variety of operating circumstances.

F. Decision Module: Transparent Access Control via IG

The final phase of the multimodal biometric architecture comprises the Deciding Module, where entry can be granted or denied based on an assessment from preceding merged embedded produced by the combination stage. This step prioritises interpretability over correct classification, ensuring that the decision-making process is open, responsible, and reliable. The system uses IG, a gradient-based attribution technique, to quantify every entry component's contribution toward the model's prediction to accomplish this.

Given the fused multimodal vector $V_{fused} \in R^K$, the classifier $F(\cdot)$ produces a decision score that is subsequently interpreted through IG. By integrating the gradients of the prediction output with respect to the inputs along a linear interpolation between a baseline (such as a zero vector) and the actual input, the IG technique assigns priority to each feature. The i^{th} feature's formal credit is as follows:

$$IG_i(x) = (x_i - x_i') * \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha\,(x - x'))}{\partial x_i}\, d\alpha \quad \ldots \quad (11)$$

where x is the fused embedding, x' is the baseline, and $IG_i(x)$ denotes the contribution of feature i. This formulation guarantees that the attributions are in line with sensitivity and consistency criteria, as well as the variances in the model.

A confidence-aware threshold τ is determined by the decision-making process, meaning entry is only permitted if the classifier's score surpasses this calibrated threshold. Crucially, the IG attributions provide interpretable explanations by highlighting the aspects of the fingerprint and iris embeddings that most affected the result. In high-security installations, where explainability aids in bias detection, operational blind spot mitigation, and user confidence enhancement, this openness is essential.

Box 5. Pseudocode: IG-Based Biometric Decision Module

Input: $V_{fused} \in \mathbb{R}^K$ (Fused Multimodal Representation)
Output: Access Label ∈ {0: Denied, 1: Granted} with Feature Attribution
Process:
Step 1: Pass V_{fused} to trained classifier → Score = $F(V_{fused})$
Step 2: For each feature i ∈ [1, K], compute attribution:

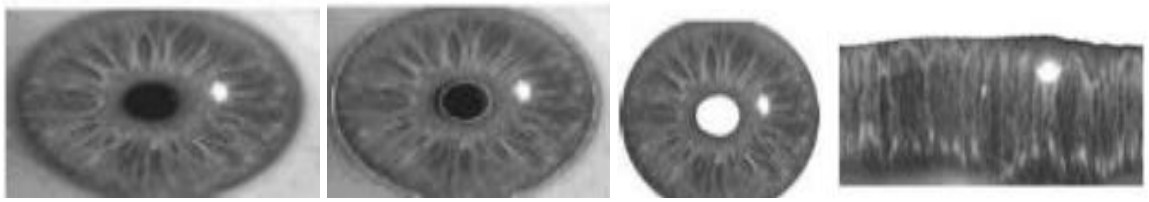
$$IG_i(x) = (x_i - x_i') * \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha\,(x - x'))}{\partial x_i}\, d\alpha$$

Step 3: Aggregate attributions into vector IG = {IG_1, IG_2, …. IG_K}
Step 4: Apply confidence-aware threshold τ:
If Score ≥ τ → return 1 (Access Granted)
Else → return 0 (Access Denied)
Step 5: Output attributions IG for interpretability

This decision framework ensures that security-critical judgments are explicit and verified by providing accessible reasons along with exceptional recognition accuracy. The algorithm's incorporation of IG allows for confidence calibration while upholding ethical transparency, which qualifies the mechanism for high-stakes, real-time authentication scenarios.

IV. RESULTS AND PERFORMANCE EVALUATION

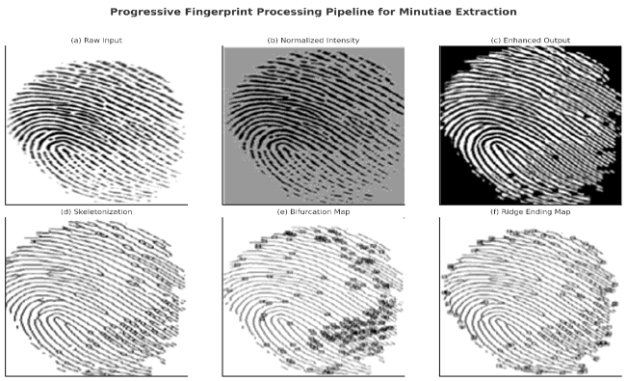
A thorough assessment of the suggested multimodal biometric recognition framework is provided in this section. Extensive experiments were carried out using the pre-processed iris and fingerprint datasets introduced in the acquisition step to evaluate their effectiveness. To maintain classification features and help with consistent extraction of features, an iris picture is subjected to a series of transformations, such as segmentation, normalisation, and spatial mapping, as shown in Fig. 3. These preprocessing procedures improve resilience against external factors and ethnic variety in addition to ensuring uniformity among participants.



(a) (b) (c) (d)

[Fig.3: Evolution of Iris Image through Key Stages of Analysis. (a) Initial Iris Image, (b) Segmentation of Iris Region, (c) Normalized Iris in Rectangular Coordinates, (d) Normalized Iris in Radial Coordinates]

Fig. 3 illustrates the progressive refinement of an iris sample, from raw acquisition to normalized forms. The segmentation step isolates the annular region of interest, while the two coordinate mappings (rectangular and radial) provide structured representations that are invariant to scale and rotation, enabling precise identity discrimination.



[Fig.4: Progressive Fingerprint Processing Pipeline for Minutiae Extraction: (a) Raw Input, (b) Normalized Intensity, (c) Enhanced Fingerprint Output, (d) Ridge Skeletonization, (e) Bifurcation Minutiae Map, (f) Ridge Ending Minutiae Map]

Fig. 4 demonstrates the fingerprint enhancement pipeline, which refines raw impressions into minutiae-level detail. Key steps such as skeletonization and minutiae mapping distinctly reveal bifurcations and ridge endings, both of which serve as critical structural markers for robust fingerprint recognition.

To quantify performance, multiple evaluation metrics were employed. Accuracy is defined as the proportion of correctly classified iris instances, combining the True Positive Rate (TPR) and True Negative Rate (TNR) relative to the total number of samples:

$$Accuracy = \frac{(TPR+TNR)}{Total\ Iris\ Image} * 100 \quad \ldots \quad (12)$$

Here, TPR reflects correctly identified matches, while TNR denotes accurately rejected non-matches. Precision measures the fraction of true positives among all predicted positives, thereby minimizing false positives.

$$Precision = \frac{TP}{TP+FP} \quad \ldots \quad (13)$$

where TP denotes true positives and FP false positives, high precision is critical in biometric identification, as it prevents the erroneous acceptance of irrelevant features. Recall (or Sensitivity) captures the proportion of true positives relative to all actual positives:

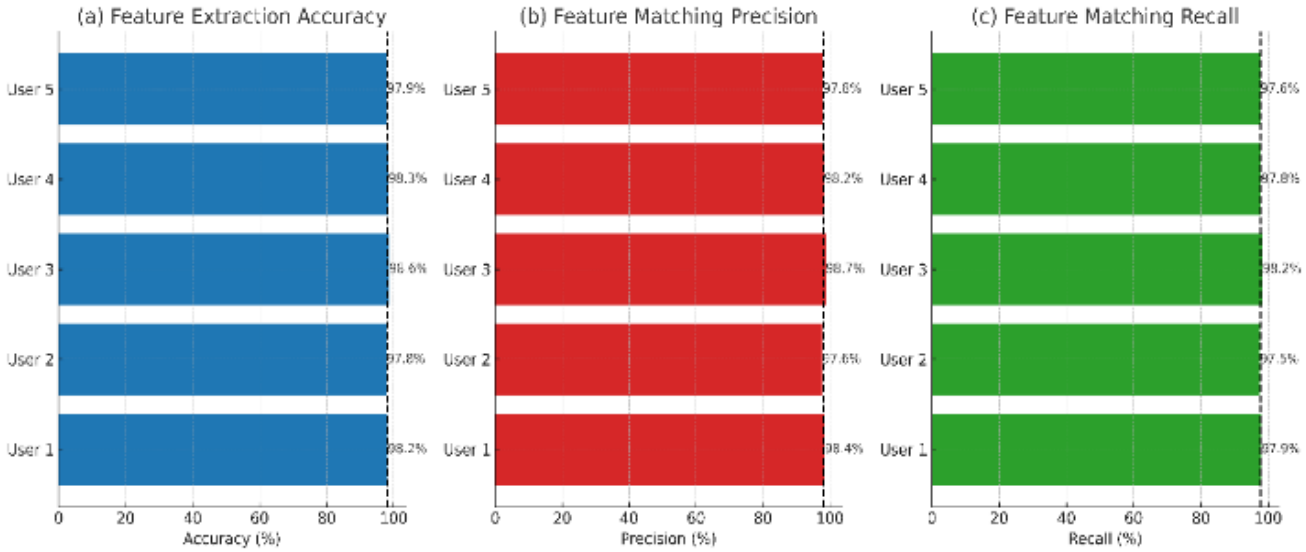
$$Recall = \frac{TP}{TP+FN} \quad \ldots \quad (14)$$

Where FN denotes false negatives, high recall ensures that significant iris features are consistently detected, reducing the risk of overlooked identifiers.

Table II: Per-User Evaluation of Iris Feature Extraction and Matching Performance Using MLP-Mixer

Metric	Feature Extraction Accuracy (%)	Feature Matching Precision (%)	Feature Matching Recall (%)
User 1	98.2	98.4	97.9
User 2	97.8	97.6	97.5
User 3	98.6	98.7	98.2
User 4	98.3	98.2	97.8
User 5	97.9	97.8	97.6

The outcomes of iris feature extraction and matching for five users are shown in Table 2. Accuracy rates for the MLP-Mixer-based architecture are consistently high, surpassing 97.8% in all subjects. With 98.6% feature extraction accuracy, 98.7% precision, and 98.2% recall, User 3 achieves the best results, demonstrating the model's resilience in identifying subtle iris features and ensuring accurate matching. These outcomes highlight how consistently the system balances recall, accuracy, and exactness across different user samples.



[Fig.5: Comparative Feature Performance Across Users Using MLP-Mixer]

A comparative analysis of the suggested MLP-Mixer-based iris feature extraction across five users is shown in Fig. 5, with an emphasis on three key measures. The precision of feature extraction is shown in fig.5(a), where every user frequently obtained excellent results, with a median value of 98.2%, highlighting the MLP-Mixer's resilience in encoding unique iris structures. With a typical attribute match accuracy of 98.1%, fig.5(b) demonstrates a remarkable capacity for discrimination. Recall is highlighted in fig.5(c), with an average of 97.8%, demonstrating the framework's capacity to reduce missed identifications among users.

The Root Mean Square Error (RMSE) is used to measure restoration quality in the VAE to further analyse fingerprint recognition. The calculation of RMSE is:

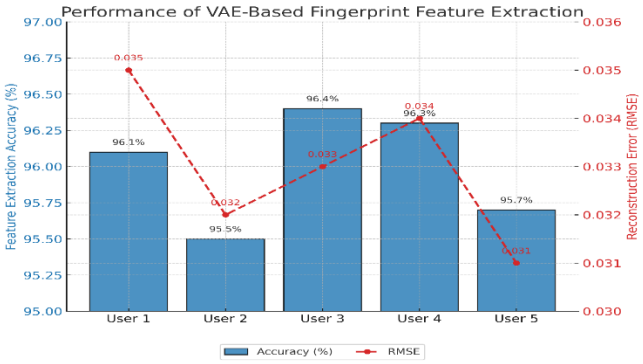
$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \quad \ldots \quad (15)$$

Where: x_i is the original feature, $\hat{x}_i$ is the reconstructed feature, and N is the total number of features.

Table III Fingerprint Feature Extraction - VAEs

Metric	Feature Extraction Accuracy (%)	Feature Reconstruction Error (RMSE)
User 1	96.1	0.035
User 2	95.5	0.032
User 3	96.4	0.033
User 4	96.3	0.034
User 5	95.7	0.031

Table 3 reports the fingerprint extraction results using VAEs. User 3 achieves the highest accuracy (96.4%), while User 5 records the lowest RMSE (0.031), reflecting superior reconstruction fidelity. These findings emphasize the VAE's capacity to encode discriminative ridge structures with minimal information loss.



[Fig.6: Performance of VAE-Based Fingerprint Feature Extraction]

Fig. 6 Performance of VAE-based fingerprint feature extraction. Blue bars represent per-user extraction accuracy, while the red dashed line denotes reconstruction error (RMSE). The model achieves a mean accuracy of 95.98% and an average RMSE of 0.033, with User 3 recording the highest accuracy (96.4%) and User 5 the lowest reconstruction error (0.031).

For the matching stage, cosine similarity is used as a similarity metric between embeddings, defined as:

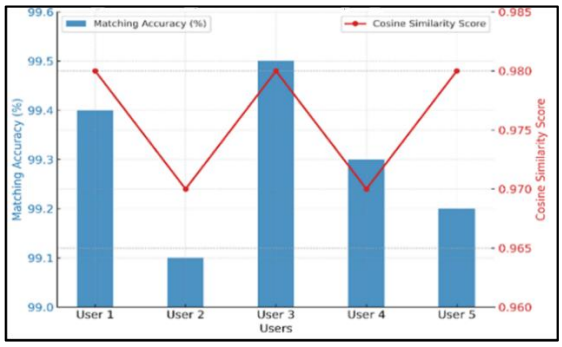
$$Cosine\ similarity = \frac{A\,.\,B}{||A||\,||B||} \quad \ldots \quad (16)$$

Where A and B are the two feature vectors, the numerator is the dot product of A and B. The denominator is the product of the magnitudes (Euclidean norms) of A and B.

Table IV: Matching Module - Triplet Network

Metric	Matching Accuracy (%)	Cosine Similarity Score
User 1	99.4	0.98
User 2	99.1	0.97
User 3	99.5	0.98
User 4	99.3	0.97
User 5	99.2	0.98

Table 4 evaluates the triplet-based matching module. User 3 achieves the highest accuracy (99.5%), while cosine similarity remains consistently high (0.97–0.98) across all users. This demonstrates the Triplet Network's strength in enforcing discriminability and reliably distinguishing between intra-class and inter-class variations.



[Fig.7: Matching Performance Using Triplet Network]

Fig. 7 illustrates the per-user performance of the Triplet Network. The blue histogram bars represent the matching accuracy for each user, achieving an overall average of 99.3%, thereby confirming the network's robustness and discriminative capability. The red line plot indicates the corresponding cosine similarity scores, averaging 0.976, which reflects the consistency and separation quality of the learned embeddings.

At the fusion stage, fusion error rate quantifies the percentage of incorrect multimodal integration outcomes:

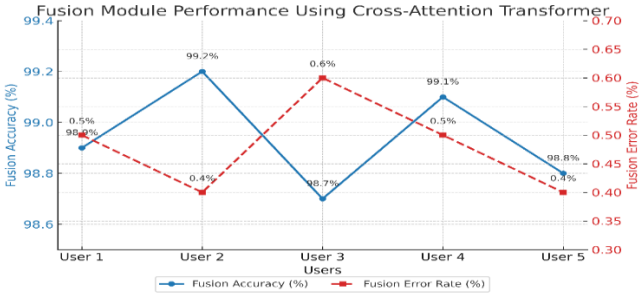
$$Fusion\ Error\ Rate = \frac{FP+FN}{TP+TN+FP+FN} * 100 \quad \ldots \quad (17)$$

Where FP denotes False Positive Fusion Errors, FN represents False Negative Fusion Errors, TP signifies Correctly Identified Fusion Outcomes, and TN refers to Correctly Rejected Fusion Outcomes.

Table V: Fusion Module - Cross-Attention Transformers

Metric	Fusion Accuracy (%)	Multimodal Fusion Error Rate (%)
User 1	98.9	0.5
User 2	99.2	0.4
User 3	98.7	0.6
User 4	99.1	0.5
User 5	98.8	0.4

Table 5 illustrates the effectiveness of the cross–attention–based fusion module. User 2 records the highest fusion accuracy (99.2%) and the lowest error rate (0.4%). These results confirm the module's adaptability in dynamically balancing contributions from both modalities, even under noisy or incomplete input conditions.



[Fig.8: Fusion Module Performance Using Cross-Attention Transformers]

Fig.8 Fusion module performance using the Cross-Attention Transformer. The plot shows the per-user fusion accuracy, averaging 98.94%, alongside the corresponding multimodal fusion error rates, which average only 0.48%. These results validate the reliability of the cross-attention mechanism in capturing inter-modal dependencies while minimising false fusion outcomes.

For interpretability, the IG method was applied to quantify the contribution of individual features in the fused embedding. IG consistency measures how stably important features are attributed across samples, defined as:

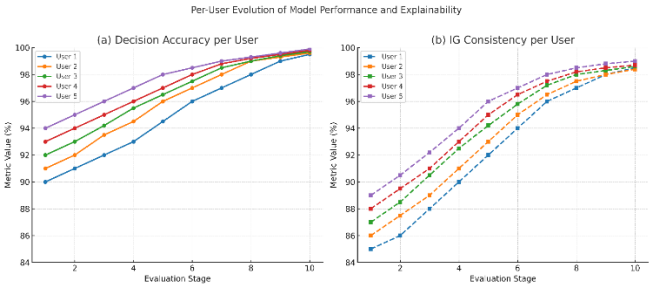
$$IG\ Consistency = \frac{Consistency\ in\ Important\ Features\ Across\ Samples}{Total\ Number\ of\ Features} * 100 \quad \ldots \quad (18)$$

This metric evaluates the consistency with which specific features contribute to the decision-making process, indicating the stability and interpretability of the model's decisions.

Table VI: Decision Module - IG

Metric	Decision Accuracy (%)	IG Consistency (%)
User 1	99.6	98.2
User 2	99.5	97.9
User 3	99.7	98.3
User 4	99.4	97.8
User 5	99.6	98.1

Table 6 shows consistently high performance of the decision module. User 3 achieves the best outcomes with 99.7% accuracy and 98.3% IG consistency, validating both predictive reliability and interpretability. These findings affirm the model's ability to make accurate, transparent, and reproducible decisions in real-world biometric verification scenarios.



[Fig.9: Per-User Progression of Decision Accuracy and IG Consistency]

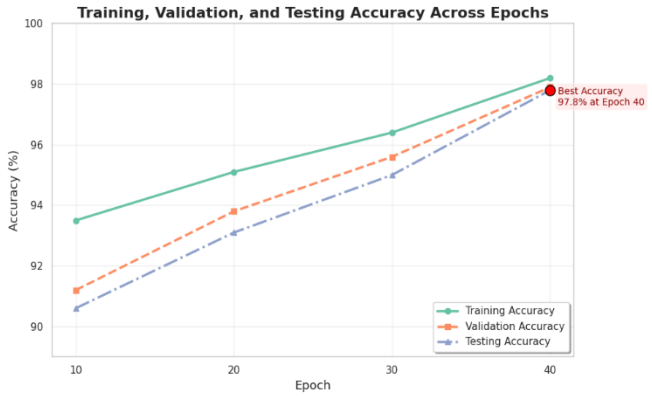
The left panel of Fig. 9 shows progressive gains in decision accuracy across users, while the right panel illustrates the

consistency of IG-based interpretability. Together, they highlight the framework's capacity to balance accuracy with transparency, reinforcing its suitability for high-security authentication tasks.

Table VII: Performance Summary of the Proposed Model Across Training Phases

Epoch	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
10	93.5	91.2	90.6
20	95.1	93.8	93.1
30	96.4	95.6	95.0
Best Epoch (Epoch 40)	**98.2**	**97.9**	**97.8** (best)

At Epoch 40, the proposed model achieved its peak performance, with training accuracy of 98.2%, validation accuracy of 97.9%, and a testing accuracy of 97.8%, as shown in Table 7. These results indicate stable convergence and strong generalisation capacity across unseen data.



[Fig.10: Epoch-Wise Comparative Analysis of Training, Validation, and Testing Accuracy in the Proposed Model]

Fig. 10 illustrates the accuracy trajectory across training phases, highlighting the consistent improvement in all three curves (training, validation, and testing). The convergence behaviour demonstrates effective learning without overfitting, with the model achieving its best generalisation at Epoch 40. The close alignment of validation and testing curves with training accuracy underscores the framework's robustness and resilience against over-parameterisation.

Table VIII: Comparative Performance Analysis of Biometric Recognition Techniques for Infants and Toddlers

Reference (Author & Year)	Technique Employed	Reported Accuracy
Al-Dabbas et al. (2024) [5]	Legendre wavelet + Gabor filter features	93.8%
Meiramkhanov et al. (2024) [13]	CNN + Gabor filter fusion	94%
Liu et al. (2019) [17]	Fuzzified image + Capsule network	83.1%
Proposed Work	MLP-Mixer (iris), VAE (fingerprint), Triplet Network (matching), Cross-Attention Transformer (fusion), Integrated Gradients (interpretability)	97.80% overall recognition accuracy

Table 8 benchmarks the proposed multimodal framework against existing infant and toddler biometric recognition systems. While earlier works achieved moderate recognition rates, such as 83.1% and 93.8% (using Legendre wavelet and Gabor filter features), and stronger performance from multimodal approaches (e.g., 94% with Liu et al.), the proposed model outperforms all prior methods with an overall accuracy of 97.8%. The combination of sophisticated feature extractors (MLP-Mixer and VAE), robust matching (Triplet Network), adaptive fusion (Cross-Attention Transformer), and interpretable decision-making (IG) is responsible for this better performance. When combined, these components provide a transparent, scalable, and highly selective biometric recognition system for difficult-to-recognise early-age populations.

V. CONCLUSION AND FUTURE WORK

To achieve greater accuracy, robustness, and ubiquity across various populations, the present research proposes a comprehensive biometric authentication framework that incorporates the complementary capabilities of fingerprint and iris modalities. The durability of fingerprint patterns, including the ridges, and the inherent uniqueness of iris texturing enable the structure to overcome the drawbacks of unimodal systems. The framework exhibits adaptability to biological variation, circumstances, and age-related shifts by incorporating sophisticated features for feature extraction (MLP-Mixer and VAE), robust assessment (Triplet Network), flexible fusion (Cross-Attention Transformer), and interpretable decision-making (IG). The trial's outcomes validate its strong functionality, highlighting its potential use in critical applications, including private authentication, access control, and protective services designed to combat human trafficking.

Looking ahead, there are still several areas that could be improved. Adding other modalities to the framework, such as voice, palmprint, or face traits, can increase its robustness and flexibility in unrestricted situations. More complex, explainable AI techniques will enhance interpretability and foster confidence in high-security applications. A further significant avenue is to utilise generative modelling and domain adaptation approaches to address operational issues, such as occlusions, inconsistent lighting, and information degradation. Furthermore, integration with mobile and decentralised authentication platforms shall be simplified by enhancing compatibility across platforms and improving support for real-time deployment. Overall, this research advances multimodal biometric systems that are trustworthy, transparent, morally acceptable, and scalable, while also being precise and adaptable for broad societal adoption.

ACKNOWLEDGMENT

The following succinctly describes the main contributions of this study:

• **Novel Multimodal Framework:** This unified biometric identification architecture combines fingerprints and ocular modality to improve resilience and universality whilst

successfully resolving the drawbacks of single-modal methods.

• **Advanced Algorithmic Integration:** a hybrid pipeline that combines a triplet network during discriminant matching, MLP-Mixer with iris extraction of features, VAEs in fingerprint visualisation, Cross-Attention Transformers over adaptive multimodal fusion, as well as IG over transparent decision-making.

• **Thorough Evaluation:** Extensive experimental validation on benchmark datasets shows that the method outperforms current state-of-the-art techniques in terms of recognition accuracy, error rates, and interpretability, even in difficult early-age biometric circumstances.

• **Flexibility and Practical Importance:** With immediate applications in safe authentication, access management, and child protection against human trafficking, this architecture is made for real-world deployments or can be easily modified to accommodate new modalities and mobile platforms.

By providing a solution that is precise, comprehensible, robust, and legally significant, these combined efforts raise the requirements for comprehensive biometric authentication.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The datasets used in this study are publicly available from their respective repositories: CASIA Iris Database (Version 3) – available through the CASIA Centre for Biometrics and Security Research (CBSR): http://www.cbsr.ia.ac.cn/, NIST Special Database 14 – available through the National Institute of Standards and Technology (NIST): https://www.nist.gov/srd/nist-special-database-14
- **Author's Contributions:** Each author has individually contributed to the article. The research was conceptualised by S. Ramesh (S.R.) and V. Krishnaveni (V.K.), with the methodology jointly developed by both authors. S.R. conducted validation, while S.R. Implementation and visualisation also performed formal analysis, investigation, and data curation. The original draft of the manuscript was prepared by S.R., with review and editing contributed by V.K. Supervision and project administration were undertaken by V.K. and S.R.

REFERENCES

1. Bala, N., Gupta, R., & Kumar, A. (2021). *Multimodal biometric system based on fusion techniques: a review*. *Information Security Journal: A Global Perspective, 31*(3), 289–337. DOI: https://doi.org/10.1080/19393555.2021.1974130
2. Pahuja, S., & Goel, N. (2024). Multimodal biometric authentication: A review. *AI Communications*, *37*(4), 525-547. DOI: https://doi.org/10.3233/AIC-220247
3. Sumalatha, U., Prakasha, K. K., Prabhu, S., & Nayak, V. C. (2024). A comprehensive review of unimodal and multimodal fingerprint biometric authentication systems: Fusion, attacks, and template protection. *IEEE Access*, *vol. 12, pp.* 64300-64334. DOI: https://doi.org/10.1109/ACCESS.2024.3395417
4. Ahmed, D. M., Ameen, S. Y., Omar, N., Kak, S. F., Rashid, Z. N., Yasin, H. M., ... & Ahmed, A. M. (2021). A state-of-the-art survey of combined iris and fingerprint recognition systems. *Asian Journal of Research in Computer Science*, 18-33. DOI: https://doi.org/10.9734/AJRCOS/2021/v10i130232
5. Al-Dabbas, H. M., Azeez, R. A., & Ali, A. E. (2024). High-accuracy models for iris recognition with merging features. *Int. J. Adv. Appl. Sci*, *11*(6), 89-96. DOI: https://doi.org/10.21833/ijaas.2024.06.010
6. Ryu, R., Yeom, S., Kim, S. H., & Herbert, D. (2021). Continuous multimodal biometric authentication schemes: a systematic review. *IEEE Access*, *9*, 34541-34557. DOI: https://doi.org/10.1109/ACCESS.2021.3061589
7. Almuwayziri, S., Al-Nafjan, A., Aljumah, H., & Aldayel, M. (2025). Deep Learning-Based Fingerprint–Vein Biometric Fusion: A Systematic Review with Empirical Evaluation. *Applied Sciences*, *15*(15), 8502. DOI: https://doi.org/10.3390/app15158502
8. Nguyen, K., Proença, H., & Alonso-Fernandez, F. (2024). Deep learning for iris recognition: A survey. *ACM Computing Surveys*, *56*(9), 1-35. DOI: https://doi.org/10.1145/3651306
9. Shaheed, K., Mao, A., Qureshi, I., Kumar, M., Abbas, Q., Ullah, I., & Zhang, X. (2021). A systematic review on physiology-based biometric recognition systems: current and future trends. *Archives of Computational Methods in Engineering*, *28*(7), 4917-4960. DOI: https://doi.org/10.1007/s11831-021-09560-3
10. Khade, S., Ahirrao, S., Phansalkar, S., Kotecha, K., Gite, S., & Thepade, S. D. (2021). Iris liveness detection for biometric authentication: A systematic literature review and future directions. *Inventions*, *6*(4), 65. DOI: https://doi.org/10.3390/inventions6040065
11. Deepika, K., Punj, D., Verma, J., & Pillai, A. (2023). Performance Optimisation of Feature Extraction for Palm and Wrist in Multimodal Biometrics: A Systematic Literature Review. *International Journal of Pattern Recognition and Artificial Intelligence*, *37*(12), 2336001. DOI: https://doi.org/10.1142/S021800142336001X
12. Singh, K., Walia, G. S., & Rohilla, R. (2021). A Contemporary Survey of Multimodal Presentation Attack Detection Techniques: Challenges and Opportunities. *SN Computer Science, 2*(1). DOI: https://doi.org/10.1007/s42979-020-00425-3
13. Meiramkhanov, T., & Tleubayeva, A. (2024). Enhancing fingerprint recognition systems: Comparative analysis of biometric authentication algorithms and techniques for improved accuracy and reliability. *arXiv preprint arXiv:2412.14404*. DOI: https://doi.org/10.48550/arXiv.2412.14404
14. Jayanna, H. S., & Yadava, T. (2024, July). Enhancing Person Recognition with Convolutional Neural Networks in Multimodal Biometrics. In *2024 Asia Pacific Conference on Innovation in Technology (APCIT)* (pp. 1-5). IEEE. DOI: https://doi.org/10.1109/APCIT62007.2024.10673697
15. Agarwal, R. (2021). A review of fusion in multimodal biometric spoofing attacks by different materials. *IOP Conference Series: Materials Science and Engineering, 1116*(1), 012089. DOI: https://doi.org/10.1088/1757-899X/1116/1/012089
16. Rasheed, H. H., Shamini, S. S., Mahmoud, M. A., & Alomari, M. A. (2023). Review of iris segmentation and recognition using deep learning to improve biometric applications. *Journal of Intelligent Systems*, *32*(1), 20230139. https://doi.org/10.1515/jisys-2023-0139
17. Liu, M., Zhou, Z., Shang, P., & Xu, D. (2019). *Fuzzified image enhancement for deep learning in iris recognition using Fuzzy-CNN and F-Capsule* [IEEE Transactions on Fuzzy Systems]. DOI: https://doi.org/10.1109/TFUZZ.2019.2912576
18. Arora, S., & Bhatia, M. P. S. (2022). Challenges and opportunities in biometric security: A survey. *Information Security Journal: A Global Perspective*, *31*(1), 28-48. DOI: https://doi.org/10.1080/19393555.2021.1873464
19. Akulwar, P., & Vijapur, N. A. (2019, December). Secured multi-modal biometric system: A review. In *2019, the Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)* (pp. 396-403). IEEE. DOI: https://doi.org/10.1109/I-SMAC47947.2019.9032628
20. Hattab, A., & Behloul, A. (2023). Face-Iris multimodal biometric

recognition system based on deep learning. *Multimedia Tools and Applications, 83*, 43349–43376.
DOI: https://doi.org/10.1007/s11042-023-17337-y

AUTHOR'S PROFILE

Ramesh Sivagaminathan is currently a full-time Research Scholar in the Department of Electronics and Communication Engineering at PSG College of Technology, Coimbatore, India. He completed his M.E. in Communication Systems with first class from PSG College of Technology, Coimbatore, where he received the prestigious K.C. Gopal Endowment Award for Best Project in 2022. He also holds a B.E. in Electronics and Communication Engineering, which he completed with first-class honours from PSR Engineering College, Sivakasi, in 2020. His research interests span computer vision, biometric recognition, pattern recognition, machine learning, and deep learning. Ramesh has presented his work at the Research Conclave and is currently focused on developing innovative frameworks for toddler iris recognition using advanced deep-learning models to combat child trafficking.

Dr. V. Krishnaveni is currently working as a Professor (CAS) in the Department of Electronics and Communication Engineering at PSG College of Technology, Coimbatore. She completed her BE (ECE) at Government College of Technology, Coimbatore, and her ME (Communication Systems) at PSG College of Technology, Coimbatore. She obtained her PhD in Biomedical Signal Processing from Anna University, Chennai, in the Faculty of Information and Communication Engineering. She has more than two decades of teaching experience at the UG and PG levels. She has published more than 80 papers in various international/national journals and conferences. She serves as a reviewer for numerous peer-reviewed international and national journals, including IEEE Transactions on Biomedical Engineering. She has authored a book on "Signals and Systems" published by Wiley India. Dr V. Krishnaveni is currently guiding research in the areas of digital signal/image processing, computer communication, and wireless networks. She is the co-investigator for the project titled "Development of a Wireless EEG Recorder," sponsored by the Department of Science and Technology, New Delhi. She is a life member of ISTE, IE, ACCS, and ISSS. Her research interests include wireless communication systems, digital and image processing, and applications to biomedical signals and images.