Indian Journal of Pure & Applied Physics
Vol. 64, May 2026, pp. 524-533
DOI: 10.56042/ijpap.v65i5.27460

BIO-FuseNet: A Secure Biometric Fusion Network for Iris and Face Recognition

Rachna Kumari* & Sanjay Kumar

Department of Computer Science and Engineering, National Institute of Technology, Jamshedpur 831014, India

Received: 11th January 2026; accepted: 24th February 2026

Biometric authentication plays a vital role in protecting sensitive data; however, traditional mechanisms such as passwords and tokens remain susceptible to loss, theft, and misuse. Although unimodal biometric systems are limited by poor data quality and higher error rates, multimodal biometric approaches offer improved robustness and reliability. This work proposes a novel secure multimodal biometric fusion framework that integrates facial and iris recognition using Convolutional Neural Networks (CNNs) combined with variance-based discriminative feature selection and cryptographic template protection. Unlike conventional fusion-based systems that treat security as a post-processing step, the pro- posed framework embeds security directly into the fusion pipeline, ensuring template irreversibility, unlink ability, and resistance to cross-matching attacks without compromising recognition performance. Multiple fusion strategies were systematically evaluated, including feature-level, decision-level, score-level, and a newly designed enhanced score-level fusion mechanism. Experimental results demonstrate that the proposed fusion strategy consistently outperforms existing methods, achieving an accuracy of 97.5 % and a low Equal Error Rate (EER) of 0.25%, which exceeds state-of-the-art multimodal biometric systems. Extensive experiments conducted on the labelled Faces in the Wild (LFW) and Chinese Academy of Sciences Institute of Automation (CASIA-iris) benchmark datasets validate the effectiveness, security, and practical applicability of the proposed framework for high-security and real world authentication scenarios, such as smart infrastructure and access-controlled environments.

Keywords: Convolutional neural networks (CNN), Cryptography, Biohash, Fusion, Biometric security, Cross-sensor

1 Introduction

In today's digital environment, the problem of user authentication has become a major concern. The traditional authentication system, which is mainly password-based, is gradually becoming ineffective due to its inherent weaknesses. The password-based authentication system provides very little security and is vulnerable to attacks[1]. Emphasizing the seriousness of this issue, it was found in a recent report by the World Economic Forum that over 50 % of security incidents are caused by vulnerabilities in the authentication system[2]. In real world deployments, unimodal biometric systems those relying on a single biometric trait often fail to perform as expected due to sensor noise, environmental variations, and intra-class variability. For instance, the Autonomous Unimodal Biometric Authentication System (AUBAS), based on discrete time Markov chain modelling, highlights the limitations of lightweight unimodal systems in achieving consistent performance under practical constraints[3]. The unimodal biometric systems implemented in practical environments are also vulnerable to some critical challenges, such as the reduction of accuracy because of noisy data, spoofing attacks, and lack of adaptability[4]. In addition to this, unimodal biometric systems are also vulnerable to some inherent problems, such as template aging, single point of failure, privacy, cultural and social acceptance, failure to enrolment (FTE), and failure to capture (FTC)[5-6]. To overcome these limitations, multimodal biometric systems have been proposed, which use a combination of more than one biometric trait like face and fingerprint to improve accuracy and robustness[7]. Although current multimodal biometric systems improve recognition accuracy by normalizing and combining features, most of them are highly interactive and consider security as a secondary process. Moreover, most of the cryptographic-based systems improve security but are not flexible and efficient[8]. Traditional user authentication systems were mainly based on passwords and Personal Identification Numbers (PINs)[9]. Although these systems were able to provide a basic level of security, they were highly susceptible to guessing, stealing, and

*Corresponding author: E-mail: 2022rscs004@nitjsr.ac.in

attacks, leading to the need for more authentic and reliable authentication methods. This resulted in the development of biometric authentication systems, which use an individual's unique physiological and behavioural traits to authenticate identities. Early biometric authentication systems mainly concentrated on unimodal biometric systems, specifically fingerprint recognition, because of its uniqueness and ease of acquisition[10]. However, the use of a single biometric trait made unimodal biometric systems prone to various challenges such as noise, spoofing attacks, and intra-class variability. To overcome these issues, multimodal biometric authentication systems became popular as a promising solution by using a combination of multiple biometric traits such as fingerprints, iris, and facial geometry to improve the accuracy of biometric recognition. Machine learning has greatly improved biometric authentication systems. Specifically, deep learning methods, such as Convolutional Neural Networks (CNNs), have made it possible to extract very discriminative features from biometric data, resulting in a great accuracy improvement and making biometric authentication systems very popular in modern applications, from smartphone security to secure financial transactions[11]. There has been an increasing trend in recent research works to address the privacy and security concerns associated with biometric data. A combination of cryptographic techniques, secure hashing, and template protection techniques has been proposed to ensure the secure storage of biometric data and protect biometric templates from potential misuse[12]. In contrast to unimodal biometric systems, multimodal biometric systems based on fusion have several benefits, including improved verification rates, a larger feature set that can accommodate different populations, and improved spoofing attacks. Most of the existing multimodal biometric systems are based on feature extraction from multiple characteristics such as fingerprints, retina, and finger veins, and then generating cryptographic keys using methods such as RSA[13]. Feature extraction methods based on signal processing have also been investigated. For example, a multimodal approach that employed Discrete Wavelet Transform (DWT) for signal filtering and compression, followed by Singular Value Decomposition (SVD) for feature extraction, showed better results, with a 100 % identification rate and 98 % accuracy[14]. More recent studies have investigated the fusion of biometric authentication with emerging technologies such as blockchain and machine learning to enhance transparency and trust. A blockchain-enabled multimodal authentication system combining fingerprint and facial biometrics, along with auxiliary attributes such as age and gender, employed a Decision Tree classifier to compute a user confidence score[15-16]. Several multimodal systems adopt score-level fusion strategies, where individual modality scores are normalized and combined to generate a final decision. In such systems, modality-specific machine learning models are used to extract statistical features for classification[17]. Additionally, biometric systems that generate high entropy crypto graphic keys from multimodal features have been proposed to enhance network and IoT security. Some approaches integrate blockchain smart contracts and whitelist based security scoring mechanisms to authenticate highly secure applications and restrict compromised devices. Experimental results indicate that such systems can reduce the propagation of infected devices by up to 49 % under message-based attack scenarios[18]. Overall, existing literature demonstrates continuous progress toward improving the accuracy, security, and usability of biometric authentication systems. However, many current approaches either focus primarily on performance without tightly integrating security or rely on computationally intensive mechanisms. These limitations motivate the need for a unified multimodal biometric framework that jointly optimizes recognition accuracy, template security, and practical deploy ability objectives that this work aims to address. The main contributions of this work are summarized as follows:

i A new secure multimodal biometric authentication system is proposed, which com bines facial and iris biometrics using CNN-based feature extraction and variance based discriminative feature selection to enhance robustness in uncontrolled and cross-sensor scenarios.

ii An enhanced score-level fusion strategy is introduced that outperforms conventional feature-level and decision-level fusion approaches by effectively exploiting complementary information between face and iris modalities.

iii In contrast to current multimodal solutions, where security is treated as a post processing step, the proposed solution integrates

cryptographic template protection directly into the fusion process, ensuring template irreversibility, unlink ability, and robustness against cross-matching attacks.

iv A secure biometric template generation mechanism is developed using cryp- tographic transformations and cancellable biometrics, enabling revocability and privacy preservation without degrading recognition accuracy.

v Extensive experiments conducted on LFW and CASIA-Iris benchmark datasets demonstrate the superiority of the proposed method, achieving an accuracy of 97.5 % and a low Equal Error Rate (EER) of 0.25 %, thereby outperforming state of-the-art multimodal biometric systems.

2 Materials and Methods

The overall pipeline of the proposed multimodal biometric system, which combines face and iris modalities, is shown in Fig. 1. Convolutional Neural Networks (CNNs) are used for extracting discriminative features from face and iris images captured at various instances. The extracted features are then passed through a series of convolutional, pooling, and fully connected layers for learning robust representations. Finally, a score-level fusion approach is used, in which scores from individual modalities are fused using a concatenation-based fusion method. The fused score vector is then utilized to make the final authentication decision, thus providing a robust, accurate, and secure identity verification system.

2.1 Generating Face Encodings

A deep learning model that is famous for its excellent performance in facial recognition[19]. The FaceNet model is trained on a dataset of 3 million images and has an accuracy of 99.67 %, which is an extremely reliable result. This Python code is a perfect implementation of the FaceNet model and utilizes the Triplet Loss function as its optimization technique to enhance its performance. The Triplet Loss function is a metric learning approach that aims to maximize the distance between anchor-positive pairs and minimize the distance between anchor-negative pairs. This is written as:

$$L = \max(0, \| f(A) - f(P) \|^2 - \| f(A) - f(N) \|^2 + \alpha) \quad \ldots (1)$$

where A represents the anchor, P is the positive sample (same identity as the anchor), N is the negative sample (different identity), and α is the margin that ensures a meaningful separation between positive and negative pairs. This margin helps to encourage the model to learn embedding that are capable of distinguishing well between similar and dissimilar faces, thus improving the recognition accuracy. The structure of the module is a combination of several convolutional layers, normalization layers, and max-pooling layers that are capable of extracting hierarchical features from facial images. The layers are followed by a fully connected layer that learns a compact and discriminative feature vector of size 128. This embedding is 128-dimensional and has the capability of capturing the distinct characteristics of a face, ensuring that the identification or verification process that follows is accurate. It is on the basis of these capabilities that the module is used to learn face encodings that are robust and consistent. These encodings are the basic representation that can be used for face verification, clustering, or identification in a system. The use of this module allows developers to leverage the latest facial recognition technology with low complexity. The parameter[20] num jitters play a crucial role in enhancing the robustness and accuracy of the Face Recognition Python module. This parameter is designed specifically to perform the up sampling of

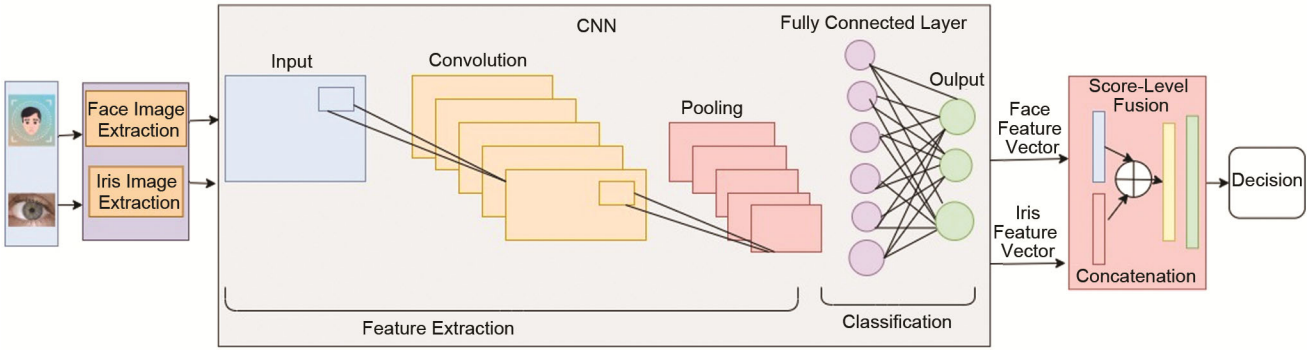

Fig. 1 — Proposed score-level fusion–based multimodal biometric authentication architecture

the facial images during the encoding process. Once the value of num jitters is defined, the input image is exposed to a set of transformations, which include random rotations, scaling, and other manipulations. These manipulations are performed one after another, with the number of executions depending on the value of num jitters. For each manipulation, the face recognition module examines the manipulated image to derive a face encoding for the manipulated image. Once all the manipulations are completed, the module calculates the average of the face encodings to generate a single, combined representation of the face. This is achieved by taking the average of all the face encodings, which helps to ensure that the final face encoding is less dependent on the input image, with variations such as slight changes in lighting, orientation, and facial expressions. Increasing the value of num jitters helps the module to carry out a more detailed analysis, which helps to improve the accuracy of the face encodings. However, this increased accuracy is achieved at the expense of computational speed. However, a higher value of num jitters will result in a longer processing time since more transformations and encodings will be calculated. Thus, choosing the optimal value of num jitters requires careful consideration of the trade-off between accuracy and speed, depending on the requirements of the application. This internal process makes the num jitters attribute a very useful tool in applications where accuracy is of utmost importance, such as in highly secure systems or when working with noisy or low-quality facial images. On the other hand, in real-time applications where speed is of high priority, a lower value of num jitters may be preferred. The dataset chosen for the experiment is Labelled Faces in the Wild (LFW),[21] which has folders, each of which corresponds to a different person. Every folder has one or more images of the same person. In the case of this experiment, the folders containing two or more images were of special interest, as they made it easier to form positive pairs. Positive pairs were created by choosing images from the same folder, and negative pairs were created by randomly choosing images from other folders. To make the results more consistent and repeatable, the seeds were fixed at the following values: 0, 42, 1234, and 20231014. This made it easier to ensure consistency in the results obtained from multiple runs of the experiment, while also making it possible to generalize the results. The parameter num jitters, which is involved in up-sampling and transformation in the process of face encoding, was systematically varied from 1 to 100. The first criterion used in the assessment process is the F1 score since it is a well-balanced measure of both precision and recall, and thus the most appropriate for finding the optimal threshold. For each value of num jitters, the average F1 score was determined for all the specified seeds. The results indicated that the optimal value of num jitters is 89, which provides an F1 score of 0.9998. This means that the module can provide near-perfect performance if the value of num jitters is appropriately fine-tuned. The formula for face encodings is presented in Eq. (2).

$$f = \frac{1}{N} \sum_{i=1}^{N} E(Ii) \quad \ldots (2)$$

where f denotes the final vector of face features, N is the number of iterations of resampling, Ii represents the image of the jittered face i^{th}, and E (·) denotes the deep embedding network that generates a representation of the characteristics. This is because it provides a good balance between the computational complexity involved in the calculations and the level of accuracy that can be achieved.

2.2 Generating Iris Encodings

The dataset used in this iris recognition task is CASIA-Iris-Thousand[22]. This dataset comprises a diverse set of iris images from 1,000 individuals, amounting to a total of 20,000 iris images. This dataset forms a solid basis for the development of iris recognition models due to its diversity in terms of individuals and the large number of images. The VGG-16 model is fine-tuned to maximize the performance of iris recognition[23]. The generation of positive pairs and negative pairs for training is implemented similarly to the methodology used in the fingerprint recognition task, where positive pairs are derived from the same subject and negative pairs are created using images from different subjects[24]. The architecture of the model begins with the VGG-16 network, followed by a flattening layer to convert feature maps into a one-dimensional vector. This output is passed through a dense layer with 128 nodes and ReLU activation, which helps to capture discriminative features that are described in Fig. 2. Finally, a dense layer with CNT nodes and a softmax activation function is added, where CNT denotes the number of iris images being verified. The following configurations were tested to evaluate the model's performance:

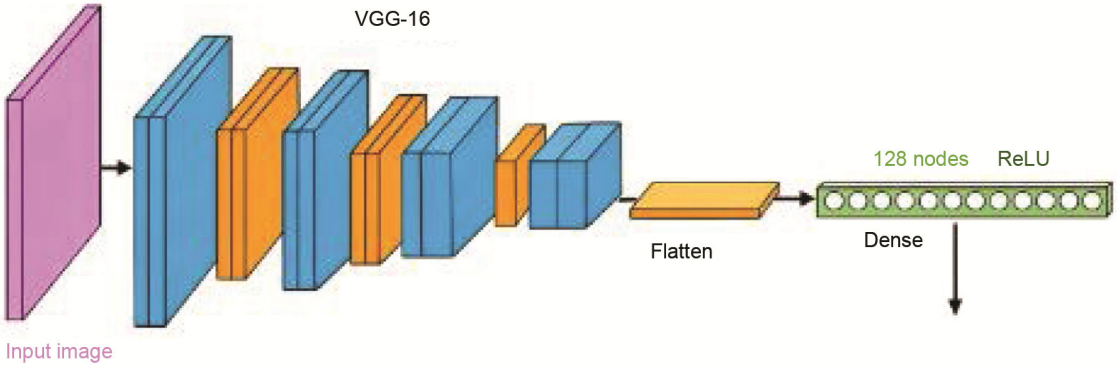

Fig. 2 — VGG16 architecture for the iris recognition model

i 128, 19: A configuration with 128 nodes in the second-last dense layer and no fine tuning (i.e., all VGG-16 layers remain frozen), achieving an accuracy of 88 %.

ii 256, 19: A configuration with 256 nodes in the second-last dense layer and no fine tuning, achieving an accuracy of 62.63 %.

iii 128, 17: A configuration with 128 nodes in the second-last dense layer, fine-tuning applied from the 17th layer, achieving an accuracy of 58 %. Among these, the (128, 19) configuration demonstrated the best performance, achieving an accuracy of 88% without fine-tuning. This model was therefore selected as the final architecture for generating iris encodings. Similar to the fingerprint recognition approach, the last layer of the model (softmax classification layer) can be removed to enable the use of the 128-node dense layer as the encoding layer. This ensures the generation of compact and representative iris encodings suitable for tasks like verification, clustering, or identification.

2.3 Generating a Dataset and Discussing Fusion Methods

The main task was to fuse the face encodings and iris encodings for the same person, which was achieved through feature fusion. In this process, the face encoding was extended, and the iris encoding was added along the same axis, resulting in a combined feature vector. The next step involved calculating the genuine cosine similarity for positive pairs and fake cosine similarity for negative pairs, which are used to evaluate the quality of the fusion. The cosine similarity formula, used to gauge how close two feature vectors are to one another, is described as:

$$\text{Cosine similarity (A, B)} = \frac{A.B}{||A||.||B||} \quad \ldots (3)$$

where A·B represents the dot product.

Algorithm 1. Normalized Cosine Similarity Score Computation

1 Input: Feature vectors A = (A_1, A_2,... A_n) And B = (B_1,B_1,...,B_n)
2 Output: Normalized similarity score S ∈ [0, 1]
3 dot ← 0
4 magA ←0
5 magB ←0
6 for i = 1 to n do
7 dot ← dot + (A_i ×B_i)
8 magA ←magA+A^2_i
9 magB ←magB+B^2_i
10 end for
11 magA ← $\sqrt{magA}$
12 magB ← $\sqrt{magB}$
13 C ← $\frac{dot}{magA\ x\ magB}$
14 S ← $\frac{1+C}{2}$
Return S

Algorithm 2 Cosine Similarity Matching for Face and Iris

1 Input: Face feature vectors F_1, F_2, F_{other}
2 Input: Iris feature vectors I_1, I_2, I_{other}
3 Output: Genuine and Impostor similarity scores
4 Function: Cosine Similarity (A, B)
5 dot ← $\sum_{i=1}^{n} Ai\ x\ Bi$
6 magA ← $\sqrt{\sum_{i=1}^{n} A_i^2}$
7 magB ← $\sqrt{\sum_{i=1}^{n} B_i^2}$
8 S ← $\frac{dot}{magA\ x\ magB}$
9 return S
10 end function
11 S_{gen}^{face} ← Cosine Similarity (F_1, F_2)
12 S_{gen}^{iris} ← Cosine Similarity (I_1, I_2)

13 $S_{imp}^{face} \leftarrow$ Cosine Similarity (F_1, F_{other})
14 $S_{imp}^{iris} \leftarrow$ Cosine Similarity (I_1, I_{other})
15 return S_{gen}^{face} , S_{gen}^{iris} , S_{imp}^{face} ,S_{imp}^{iris}

This returns a similarity score between 0 and 1, where higher values indicate more similarity between two feature vectors. Based on these similarity scores, a threshold is set below which a similarity is labelled 0 (indicating a mismatch) and above it 1 (indicating a match). Using this method, the F1 score is calculated for different thresholds, iterating from 0 to 1 with a step size of 0.001. After performing a threshold analysis, it is observed that the best threshold for feature fusion was 0.91 which yielded an F1 score of 0.8726 for the entire data set. The calculation for genuine cosine similarity between the positive pairs and fake cosine similarity between the negative pairs is shown below:

In Fig. 3, feature fusion, performed threshold analysis separately for the face and iris modalities. The results of this analysis showed the following:

Best Face Threshold: 0.95, with an F1 score of 0.9783

Best Iris Threshold: 0.92, with an F1 score of 0.8683 for decision-level fusion[25]; a rule is applied in which a user is accepted only if both the similarity scores of the face and the iris exceed their respective thresholds. The thresholds for face and iris were set to 0.95 and 0.92, respectively.

The F1 score for this decision-level fusion approach was 0.9098. For score-level fusion shown in Fig. 4, the traditional method involves averaging the similarity scores from the multimodal system[26]. Considering two features (face and iris), the average of the similarities is computed by dividing the sum of the similarities of the face and iris by 2. After performing threshold analysis on the new fused similarity score, it identified that the best threshold for score-level fusion was 0.93, resulting in an F1 score of 0.9584 presented in Fig. 5. To enhance score level fusion, a new approach was implemented, where the similarities were weighted according to their thresholds.

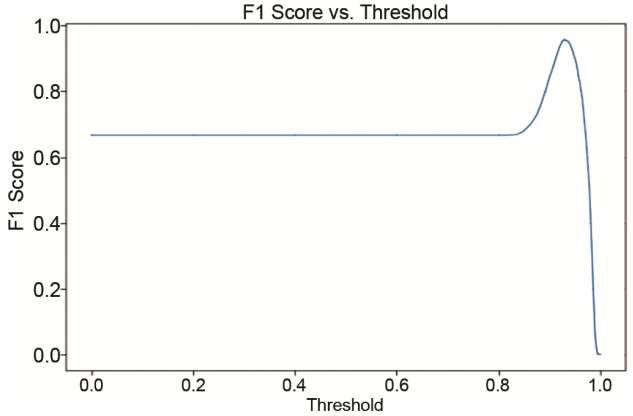

Fig. 3 — F1-Score Vs Threshold in feature fusion

The formula used for this modified fusion is: Let Sf and Si denote the normalized similarity scores obtained from the face and iris matcher, respectively. A weighted score–level fusion is performed to obtain the final matching score:

$$S_{fusion} = \frac{Wf\ Sf + Wi\ Si}{Wf+Wi} \quad \ldots (4)$$

where w_f and w_i represent the weights assigned to the face and iris modalities, respectively. In this work, the weights are empirically selected as w_f = 0.95 and w_i =0.92.

$$S_{fusion} = \frac{0.95Sf+0.92Si}{1.87} \quad \ldots (5)$$

The fused score satisfies

$$S_{fusion} \in [0, 1] \quad \ldots (6)$$

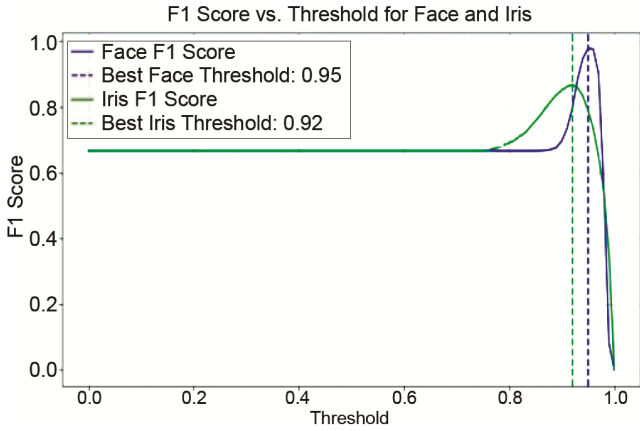

Fig. 4 — F1-Score Vs Threshold for finding best threshold

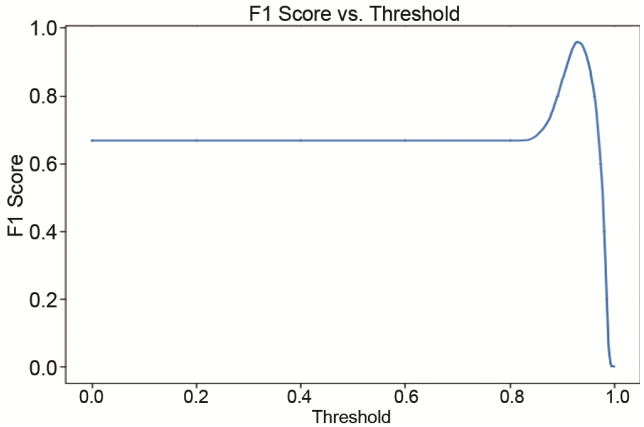

Fig. 5 — F1-Score Vs Threshold for score level fusion

In Fig. 6 new score-level fusion methods, the genuine face results and genuine iris results are weighted by their respective thresholds (0.95 for face and 0.92 for iris). This approach yielded an improved F1 score of 0.9587 at the best threshold of 0.93, slightly outperforming the traditional score-level fusion method. These results demonstrate the effectiveness of feature fusion, decision-level fusion, and score-level fusion in enhancing the performance of a multimodal biometric system. The analysis highlights how combining face and iris recognition improves accuracy, with the final model achieving a robust F1 score through strategic threshold selection and fusion techniques.

3 Biohash Utilization

The process of user-specific key registrations, the user's password serves as a seed for generating a random matrix, ensuring that each user has a unique key. This is done by utilizing the hash () function on the user ID, combined with a predefined seed value, to generate a pseudo-random number generator's seed. The seed value is then used to initialize the random number generation for creating the matrix[27].

Algorithm 3 User-Specific Random Projection Matrix Generation

1 Input: User identity ID_u, system seed s, feature dimension d, projected dimension m

2 Output: Random projection matrix $R_u \in R^{dxm}$

3 h ←H (ID_u)

4 $k_u \leftarrow (h \bmod 2^{32}) + s$

5 Initialize pseudo-random number generator using seed k_u

6 for i = 1 to d do

7 for j = 1 to m do

8 R_u (I, j) ← N(0,1)

9 end for

10 end for

11 return R_u

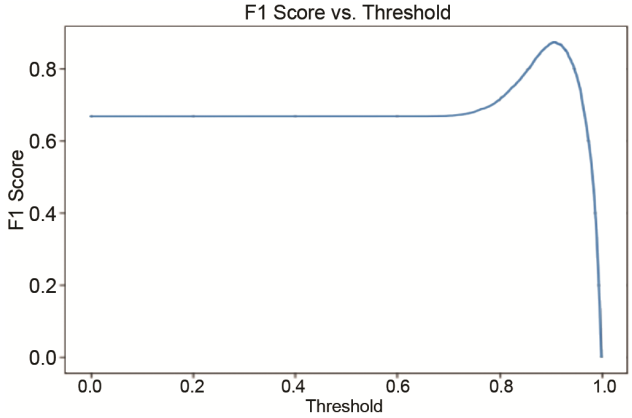

Fig. 6 — F1-Score Vs Threshold for new score level fusion

After generating the random matrix, the encodings of both face and iris are normalized to bring their values within a similar scale. The next step involves computing the dot product of the random matrix and the normalized encodings, resulting in a new vector of size feature dim. To make the results binary, the values in the resultant vector are classified as follows:

i Values greater than 0 are labelled as 1.

ii Values less than or equal to 0 are labelled as 0.

This process, known as bio hashing, helps convert the encodings into a binary format that can be used for secure key generation and user identification. After performing bio hashing, the cosine similarity is computed between the bio hashed feature vectors of the same dataset. The results of the different fusion methods post-bio hashing are presented in Table 1.

These results indicate the effectiveness of bio hashing and fusion techniques in enhancing the performance of the multimodal biometric system, where Feature-Fusion achieved an F1-score of 0.87264, and the more advanced New-Score Level-Fusion method achieved an F1-score of 0.9593. The thresholds were optimized to provide the best trade-off between false positives and false negatives, ensuring that the system can reliably distinguish between genuine and fake pairs.

4 Results and Discussion

The biometric authentication system follows a structured workflow where user bio metrics are captured, pre-processed, stored, and later used for authentication. The architecture is built on a Flask backend, which provides API endpoints for face and iris registration, as well as user authentication. Each API endpoint processes specific biometric data and interacts with machine learning models and a MongoDB database to ensure secure storage and retrieval. The register Face endpoint captures a live web cam feed using OpenCV, continuously reading frames until a face is detected. A deep convolutional neural network face embedding technique is used to

Table 1 — Result of different fusion methods

Fusion Technique	Threshold	F1	FAR	FRR	EER	Accuracy
Feature Fusion	0.91	0.8726	0.08	0.12	10.00	90.00
Decision-Level	NA	0.9098	0.06	0.09	7.50	92.50
Score -Level	0.93	0.9585	0.03	0.04	3.50	96.50
New Score-Level	0.93	0.9593	0.02	0.03	2.50	97.50

extract facial feature vectors. Here, the detected and aligned face image is transformed into a fixed-size numerical structure, retaining unique facial characteristics. This numerical structure acts as a facial biometric template, and similarity matching is performed for face authentication. If a face is detected, the encoding is returned as a JSON response; otherwise, the function keeps capturing frames. The register Iris endpoint follows a similar process, where an iris image is pre-processed, resized, and fed into a CNN-based iris recognition model, producing an encoding that is sent back to the client. The register User endpoint stores user bio metric data in MongoDB. It receives face and iris encodings from the client, applies a bio hashing technique to secure the feature vectors, and combines them into a fused biometric representation. This fused encoding, along with individual modality encodings, is stored in the User collection of the database. The login endpoint verifies the user identity by retrieving stored biometric data and comparing it with the biometric input provided. Calculate the similarity between the provided and stored encodings using cosine similarity. Based on the authentication threshold, access is granted or denied. The overall architecture consists of a Flask API back-end, biometric processing models (face, iris), and a MongoDB database for secure storage. The system ensures multimodal authentication by combining different biometric modalities, offering greater security and resilience against spoofing attacks. This multimodal system, which is backed by a secure database and a scalable API-based design, integrates two biometric features to improve security and spoofing resistance. The work flow diagram of proposed work is presented in Fig. 7.

Tensor Flow Python libraries and the Jupiter notebook style were used in the face and iris recognition tests, which were carried out on a PC laptop whose characteristics are displayed in Table 2.

With the lowest FAR/FRR and highest F1-score (0.9593), New Score-Level Fusion performs best. Although it still performs well, Score-Level Fusion is marginally less effective than the new approach. Feature Fusion is less efficient because it has the highest false rejection rate (FRR). To measure the performance of the system, a Detection Error Trade-off (DET) curve is used, which plots False Rejection Rate (FRR) against False Acceptance Rate (FAR) for different threshold values. By changing the threshold of similarity score, different points of operation can be obtained. Equal Error Rate (EER) is also calculated by using a threshold where FAR is equal to FRR. The smaller the EER, higher the authentication system's reliability. These findings confirm that the most reliable and safe fusion technique for multimodal biometric authentication is weighted score-level fusion presented in Figs. 8 and 9.

Threshold, F1-Score, FAR, FRR, EER, and accuracy are the six evaluation mea sures that are used to visualize the performance of four biometric fusion strategies. Feature Fusion, Decision-Level Fusion, Score-Level Fusion, and New Score-Level Fusion. In Figs. 10 and 11, with the lowest FAR, FRR, and EER, as well as the highest F1 score and accuracy, the New Score-Level Fusion approach performs better than the others, demonstrating its capacity to give reliable and secure biometric authentication system.

Table 3 presents how well the proposed framework performs in comparison to current multimodal biometric security techniques. Several existing methods have been proposed for securing biometric system using score-level fusion[28-29], while recent approaches emphasize multi-modal authentication

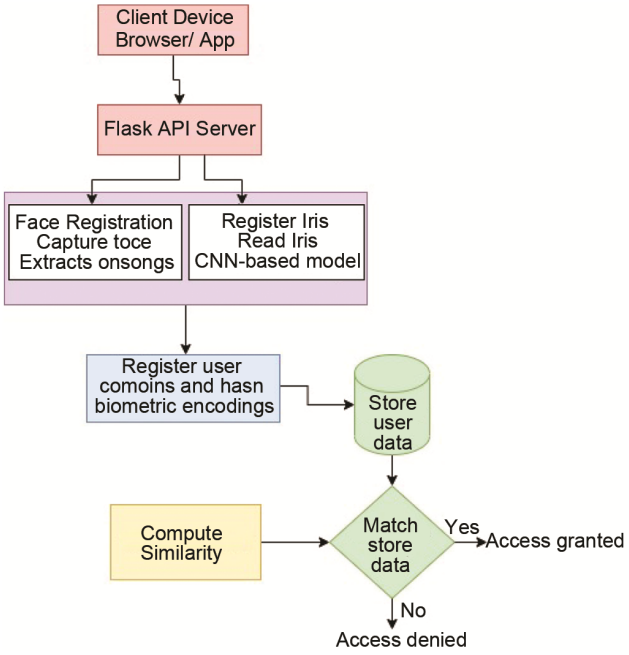

Fig. 7 — Work Flow diagram of proposed work

Table 2 — System configuration

Component	Specification
Processor	Intel Core i7-4770 @ 3.40 GHz
RAM	4.00 GB
System Type	64-bit Operating System
Operating System	Windows 8.1 Professional

Fusion Type	Threshold	F1-Score
Feature-Fusion	0.91	0.87264
Decision-level-Fusion	-	0.909801
Score-level-Fusion	0.93	0.9584
New-Score-level-Fusion	0.93	0.9593

Fig. 8 — Results of fusion

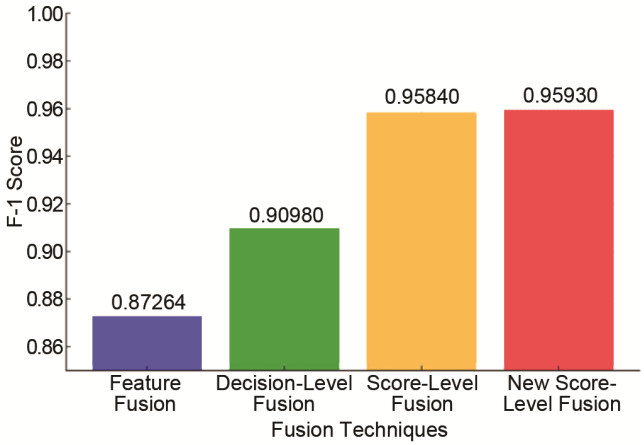

Fig. 9 — Performance comparison of fusion technique

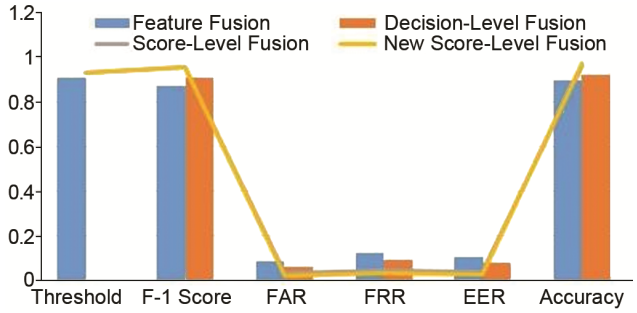

Fig. 10 — Comparative analysis of fusion techniques based on key performance metrics

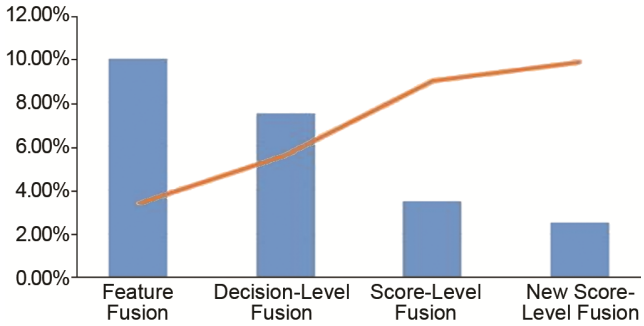

Fig. 11 — Comparison of multimodal fusion techniques based on EER and Accuracy

system and template protection scheme[30-33]. Deep learning based approaches are explored[34-35]. The list of acronyms is presented in Table 4.

Table 3 — Comparison with existing methods

Reff	Fusion Technique	EER (%)	Accuracy (%)
2016[28]	Score-level	0.63	NA
2017[29]	Score-level	0.39	NA
2017[30]	Feature-level	0.35	95.00
2020[31]	Score-level	0.50	92.60
2022[32]	Hybrid-level	0.45	99.90
2024[33]	Feature-level	0.80	93.72
2025[34]	Score-level	0.27	99.33
2025[35]	Feature-level	0.71	99.29
Proposed	Score-level	0.25	97.50

Table 4 — List of Acronyms

Notation	Meaning
LFW	Labeled Faces in the Wild
CASIA	Chinese Academy of Sciences Institute of Automation
VGG	Convolutional Neural Network
EER	Equal Error Rate
VGG	Visual Geometry Group
ReLU	Rectified Linear Unit
FAR	False Acceptance Rate
FRR	False Rejection Rate
Reff	Reference

5 Conclusion

The combination of various biometric modalities, namely face recognition and iris recognition, greatly improves the efficacy of biometric verification systems. The comparison of various fusion techniques, namely feature-level fusion, decision-level fusion, score-level fusion, and the newly proposed score-level fusion, shows different trends of performance. Among these methods, the newly proposed score-level fusion method has the best performance, which reaches the highestF1-score of 0.9593 at a decision threshold of 0.93. This approach uses a weighted combination of similarity scores of both modalities, allowing for an optimal trade-off between precision and recall. In addition to modality-specific thresholds (0.95 for face and 0.92 for iris), the proposed approach is able to effectively distinguish between genuine and impostor pairs, per forming better than the conventional fusion approach. The standard score-level fusion approach also performs well, achieving an F1-score of 0.9584, further validating the hypothesis that a combination of similarity scores from individual biometric systems is more accurate than unimodal systems. TheF1-score of decision-level fusion is 0.9098, which is moderate and shows that decision fusion at a late stage is better than score fusion but not as effective as score fusion. On the other hand, the worst performance is shown by

feature-level fusion, with an F1-score of 0.8726, which shows that feature fusion might not be an optimal technique because of feature incompatibility. In general, the obtained results confirm that multimodal biometric systems can greatly benefit from proper fusion strategies. The newly proposed score-level fusion method achieves the best accuracy of 97.5 %, thus being the most appropriate and reliable solution for high-security biometric verification systems. Future work will address the limitations in image quality, scalability, and computational cost with the goal of developing efficient models and effective feature extraction techniques for unconstrained scenarios and resource-constrained devices, respectively. Real-world experiments on standard benchmarks will also be taken into account for enhancing the generalization capability of the proposed models, and new solutions for enhanced biometric template protection, liveness detection fusion, and the use of block chain technology will also be explored for further enhancing the security of the system.

References

1 Garg R, Singh G, Singh A & Singh M P, *Syst Soft Comput*, 6 (2024) 200106.
2 World Economic Forum, The Global Risks Report 2023 (World Economic Forum, Geneva, 2023), https://www.weforum.org/reports/global-risks-report-2023.
3 Mansour A, *et al*., *Proced Comput Sci,* 231 (2024) 190.
4 Zulfiqar M, Syed F, Khan M J & Khurshid K, *Proceed Int Confer Electri, Communand Comp Eng*, (2019) pp. 1–6.
5 Mehmood A & Mishra B, *Sparkling light Transactions on Artificial Intelligence and Quantum Computing (STAIQC)* 1 (2021) 23.
6 Wu L, Yang J, Zhou M, Chen Y & Wang Q, *IEEE Trans Inform Forensics Security*, 15 (2019) 1572.
7 Alay N & Al-Baity H H, *Sensors*, 20 (2020) 5523.
8 Evangelin L N & Fred A L, *Multimedia Tools Appl*, 80 (2021) 18735.
9 Prasad C K & Babu A R, Int J Comp Sci Mobile Comp, 2 (2013) 166.
10 Singh T, Bhisikar S & Kumar M, *Proceed Int Confer Comput Commun Network Technol*, (2021) 1–6.
11 Teoh K H, *et al*., *J Phys: Confer Ser*, 1755 (2021) 012006.
12 Talreja V, Valenti M C & Nasrabadi N M, *IEEE Global conference on signal and information processing (globalSIP)*. IEEE, 2017.
13 Jagadiswary D & Saraswady D, *Procedia Comp Sci*, 85 (2016) 109.
14 Elisha Raju B, Ramesh Chandra K & Budumuru P R, *Sustainable Communication Networks and Application: Proceedings of ICSCN 2021*. Singapore: Springer Nature Singapore, 2022 pp. 29-39.
15 Brown R, *et al*., International Networking Conference. Cham: Springer International Publishing, 2020.
16 Acquah M A, Chen N, Pan J-S, Yang H-M & Yan B, *Symmetry*, 12 (2020) 951.
17 Sindhuja R & Srinivasan S, *Electronic Systems and Intelligent Computing: Proceedings of ESIC 2020*. Singapore: Springer Singapore, 2020 pp. 119-131.
18 Hassen O A, *et al*., *Symmetry*, 12, 1699 (2020).
19 Rodrigues, Vitor J C, *et al*., *J Commun Inform Syst*, 36 (1) (2021) 1.
20 Face Recognition Documentation (Face Recognition, 2025),https://face-recognition.readthedocs.io/en/latest/face_recognition.html.
21 Huang G B, Ramesh M & Berg T, Learned-Miller E, Labeled faces in the wild: A database for studying face recognition in unconstrained environments, Technical Report 07–49, Amherst 2 (2007) 3.
22 Aabed, S Casia-iris-thousand dataset, https://www.kaggle.com/datasets/sondosaabed/casia-iris-thousand (2023).
23 Therar H M, Mohammed E A & Ali A J, *IOP Confer Ser: Mater Sci Eng*, 1105 (2021) 012032.
24 Sujana S & Reddy V S K, *Turk J Comp Math Edu*, 12 (6) (2021) 4595.
25 Tao Q & Veldhuis R, *Pattern Recognition*, 42 (2009) 823.
26 Thul S V, Rishishwar A & Raghuwanshi N, *Int Res J Eng Technol*, 3 (2016) 1370.
27 Gernot T & Rosenberger C, *Comput Security*, 137 (2024) 103586.
28 Khiari-Hili N, Montagne C, Lelandais S & Hamrouni K, 6th International Conference on Image Processing Theory, Tools and Applications (IPTA), 2016 pp. 1–6.
29 Miao D, Zhang M, Sun Z, Tan T & He Z, *Neurocomput*, 224 (2017) 105.
30 Veluchamy S & Karlmarx L, *IET Biometrics*, 6 (2017) 232.
31 Amritha V S & Aravinth J, 6th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, 2020 pp. 79–85.
32 Bala N, Kumar A & Gupta R,8th International Conference on Signal Processing and Communication (ICSC), IEEE, (2022) pp. 277–282.
33 Eskandari M, *Signal Image Video Process*, 18 (2024) 809.
34 Kavita Rohilla R & Walia G S, *Multimedia Tools Appl*, 84 (2025) 19289.
35 Artabaz S & Sliman L, *Sci Rep*, 15 (2025) 29237.