IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 4, Ser. 1 (July. – August. 2026), PP 13-24
www.iosrjournals.org

Deep Feature Fusion of Iris and Retina for Robust Person Identification

Md. Anisuzzaman Siddique, Md. Morshedul Arefin, Mahboob Qaosar, Asif Zaman
{anisuzzaman, moshedul, qaosar, asif}@ru.ac.bd
Department of Computer Science and Engineering, University of Rajshahi, Bangladesh

Abstract
Reliable person identification is critical for modern security systems, yet unimodal biometric methods often suffer from noise, environmental variations, and user-related differences. To overcome these limitations, this paper presents a robust deep learning–based multimodal biometric system for person identification using deep feature fusion of iris and retina traits. Separate convolutional neural networks are used to extract highly discriminative features from iris and retinal images, capturing complementary information from each modality. The extracted features are fused at the feature level and further combined using weighted fusion strategies to enhance recognition performance. Stable iris regions and retinal blood vessel patterns not only improve accuracy but also support liveness detection to reduce spoofing attacks. Extensive experiments conducted on publicly available iris and retina datasets, evaluated using standard biometric metrics such as False Acceptance Rate (FAR), False Rejection Rate (FRR), Equal Error Rate (EER), and ROC analysis, demonstrate that the proposed approach significantly outperforms unimodal systems and conventional fusion methods. The results confirm that deep feature fusion of iris and retina provides a secure, reliable, and efficient solution for next-generation biometric authentication systems.
Keywords: *Multimodal Biometrics, Iris Recognition, Retina Recognition, Deep Learning, Feature Fusion*

Date of Submission: 01-07-2026 Date of Acceptance: 10-07-2026

I. Introduction

Reliable person identification is a critical requirement for modern security and access control systems. Traditional authentication methods, such as passwords, PINs, and ID cards, are vulnerable to theft, loss, and misuse. In contrast, biometric systems offer a more secure and convenient solution by leveraging physiological or behavioral traits that are unique to each individual, including fingerprints, facial features, voice, iris, retina, hand geometry, and signature [8]. Despite their advantages, unimodal biometric systems often face challenges such as sensor noise, environmental variations, intra-class variability, and non-universality, which can reduce recognition accuracy and limit applicability in high-security environments [3].

Multimodal biometric systems, which combine two or more biometric traits, provide a robust solution to these challenges. By integrating complementary information from multiple traits, these systems enhance recognition accuracy, reduce error rates, improve population coverage, and provide stronger protection against spoofing attacks [4]. Fusion in multimodal systems can occur at the feature, score, or decision level. Feature-level and score-level fusion are widely preferred, as they retain more discriminative information while enabling flexible and effective integration [6]. Recent advances in deep learning have significantly improved biometric recognition by enabling automatic learning of highly discriminative feature representations from biometric images. In particular, CNN-based architectures have become the dominant approach for iris recognition, segmentation, presentation attack detection, and multimodal biometric systems, demonstrating superior performance over traditional handcrafted feature extraction techniques [30], [31].

Among biometric traits, **iris and retina** are particularly reliable due to their uniqueness, stability, and resistance to forgery. Iris patterns, located in the annular region between the pupil and sclera, remain stable throughout life and contain highly discriminative texture features. Retinal blood vessel patterns are also distinctive and provide natural liveness detection, as retinal imaging is possible only from living individuals [10], [21]. These characteristics make iris and retina highly suitable for applications requiring stringent security.

Recent advances in **deep learning** have further improved biometric recognition by enabling automated extraction of discriminative and robust features from raw images. Deep neural networks outperform traditional handcrafted methods by effectively handling noise, illumination changes, and alignment issues. When combined with multimodal fusion strategies, deep learning enables highly accurate and reliable person identification even in challenging conditions [6].

Although recent deep learning approaches have significantly improved unimodal biometric recognition, limited attention has been devoted to exploiting complementary information from both iris and retinal modalities using deep feature fusion. This motivates the development of an effective multimodal framework capable of learning robust and discriminative biometric representations [30], [31]. Motivated by these advantages, this paper proposes a **deep feature fusion–based multimodal biometric system** using iris and retina for robust person identification. Features from each modality are extracted using separate deep models and then fused to exploit complementary information. The system aims to improve recognition accuracy, reduce False Acceptance Rate (FAR), False Rejection Rate (FRR), and Equal Error Rate (EER), and provide secure, liveness-aware authentication.

The main contributions of this work are summarized as follows:

- A **Deep Feature Fusion-Based Person Identification Framework** is proposed that combines iris and retinal biometrics for accurate and reliable person identification.
- A **feature-level fusion strategy** is developed to integrate deep features extracted by Iris-CNN and Retina-CNN, improving the discriminative capability of the fused representation.
- A **hybrid loss function**, combining categorical cross-entropy and triplet loss, is employed to enhance feature learning and improve recognition performance.
- The proposed framework is evaluated on benchmark datasets using standard biometric metrics, achieving **99.12% recognition accuracy** and demonstrating superior performance over unimodal and conventional fusion methods.

The remainder of this paper is organized as follows. **Section II** reviews the related work on multimodal biometric person identification. **Section III** presents the proposed Deep Feature Fusion-Based Person Identification Framework, including image preprocessing, deep feature extraction, feature fusion, the hybrid loss function, and the classification process. **Section IV** describes the experimental setup, benchmark datasets, evaluation metrics, and implementation details. **Section V** discusses the experimental results and performance analysis. Finally, **Section VI** concludes the paper and outlines directions for future research.

II. Related Work

Multimodal biometric systems have been widely explored to address the limitations of unimodal systems, which are often affected by noise, environmental variations, and intra-class differences [8], [7]. Fusion in multimodal systems can occur at the feature, score, or decision level. Score-level fusion is commonly preferred because it allows the combination of heterogeneous matching scores while retaining flexibility and robustness [8], [3].

In iris recognition, conventional systems often combine information from left and right irises using simple classifier rules [9]. Daugman [6] improved iris recognition through active contour-based localization, Fourier-based off-axis handling, eyelash detection, and score normalization. Other methods addressed iris noise removal [10, 11], probabilistic pattern matching [12], and DCT-based coding for robust feature extraction [13]. More recently, deep learning has transformed iris recognition by replacing handcrafted descriptors with automatically learned hierarchical representations. Recent survey studies have reported that CNN-based and transformer-based models provide improved robustness, higher recognition accuracy, and better generalization across challenging datasets [30], [31].

Retinal recognition has also evolved from full vessel-tree approaches [14, 15] to landmark-based extraction using bifurcations and crossovers for compact and efficient templates [16], [17]. Frequency and spatial-domain techniques have been applied for retinal feature extraction [18], [19]. Waheed et al. [24] proposed a fast, non-vascular retinal recognition system achieving over 92% accuracy on standard datasets, while Lajevardi et al. [21] used Biometric Graph Matching combined with SVM classifiers and Kernel Density Estimation for robust performance on small datasets. Recent retinal biometric research has increasingly focused on deep feature learning and graph-based vascular representation to improve recognition performance under varying imaging conditions. Public retinal datasets continue to play an important role in benchmarking and comparative evaluation of retinal recognition algorithms [32], [33].

Several studies have explored **iris-retina fusion**. Saha et al. [22] applied PCA-based feature-level fusion, achieving 98.37% recognition compared to 96.74% and 94.56% for individual modalities. Choras [23] and Modarresi et al. [24] used Gabor and Contourlet transforms to improve accuracy. Sen and Islam [25] combined log-Gabor iris features with discrete wavelet retina features, achieving FAR 2.041%, FRR 0%, and RR 97.959%. Latha et al. [26] proposed hybrid fusion of frequency-domain iris and spatial-domain retina features using weighted-score fusion, reaching ERR 0.01% and RR 99.3%. Kihal et al. [27] combined iris and corneal shape features at the score level, reducing EER to 0%.

Multimodal systems combining other traits, such as face, fingerprint, palmprint, hand geometry, and voice, have also demonstrated substantial performance gains. For instance, Hong and Jain [1] showed that PCA-based face and minutiae-based fingerprint fusion reduced FRR from 61.2% and 10.6% for unimodal systems to 6.6%. Frischholz and Dieckmann [2] and Snelick et al. [6] demonstrated similar improvements using decision- and score-level fusion.

In summary, **fusion-based multimodal systems**, particularly iris-retina combinations, consistently outperform unimodal approaches in accuracy, robustness, and spoof-resistance. However, many approaches rely on frequency-domain features or minutiae-based extraction, which increase computational costs. This motivates the development of efficient **spatial-domain deep feature fusion methods** for practical, real-time iris-retina biometric authentication [28, 29**].**

III. Enrollment And Testing In A Biometric System

Figure 1 illustrates the operational workflow of a biometric recognition system, encompassing both the enrollment and testing phases. The system is designed to acquire biometric traits from individuals, extract distinctive features, and perform either verification or identification based on stored biometric templates.

The process begins with the presentation of a biometric trait, such as an iris, retina, fingerprint, face, or other physiological characteristic. The biometric sample is captured using an appropriate sensor capable of acquiring high-quality biometric data. The acquired sample is then subjected to sampling, segmentation, and feature extraction procedures. During this stage, the region of interest is identified, relevant biometric patterns are isolated, and discriminative feature vectors are generated to represent the individual's unique biometric characteristics.

Following feature extraction, a quality control mechanism evaluates the captured sample to ensure that it satisfies predefined quality requirements. Factors such as image clarity, illumination, noise level, and segmentation accuracy are assessed. If the quality of the biometric sample is deemed inadequate, the system requests reacquisition of the biometric trait. This iterative process continues until an acceptable sample is obtained.

Once a high-quality biometric template has been generated, the system determines whether the user is undergoing enrollment or testing. During the enrollment phase, the extracted biometric features are converted into a biometric template and stored in the biometric enrollment database. The database serves as a repository of enrolled users and their associated biometric templates for future authentication or identification purposes.

During the testing phase, the system operates in either verification mode or identification mode. In verification mode, the user claims a specific identity, and the extracted biometric template is compared against the corresponding enrolled template stored in the database. This process performs a one-to-one (1:1) comparison to determine whether the presented biometric sample matches the claimed identity. A similarity score is generated and evaluated using a predefined decision rule. If the similarity score exceeds the acceptance threshold, the user is classified as genuine; otherwise, the user is classified as an impostor.

In identification mode, the user does not provide a claimed identity. Instead, the extracted biometric template is compared against all enrolled templates stored in the biometric database. This process performs a one-to-many (1:M) comparison to locate the best matching enrolled template. The resulting similarity score is evaluated using a decision rule to determine whether a valid match exists. If the similarity score satisfies the predefined acceptance criterion, the system reports that the identity has been found. Otherwise, the system concludes that no matching identity exists within the database and reports that the identity is not found.

The decision-making modules play a critical role in controlling system performance by balancing the trade-off between false acceptance and false rejection. Appropriate threshold selection enables the biometric system to achieve reliable recognition while maintaining high security and accuracy.

Overall, the biometric recognition framework integrates data acquisition, quality assessment, feature extraction, template management, matching, and decision-making into a unified architecture. The enrollment process establishes the biometric database, whereas the testing process performs verification or identification by comparing query templates against enrolled templates and generating authentication decisions based on similarity measures.

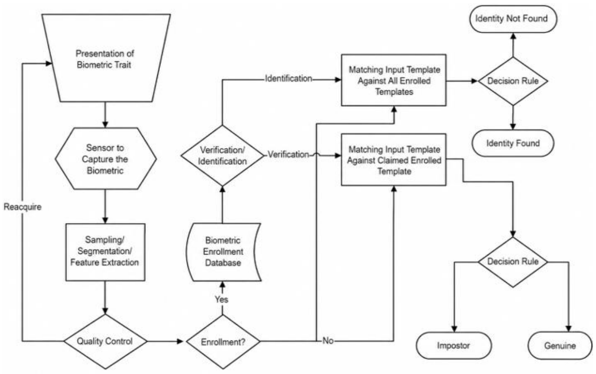

Figure 1: Enrollment and Testing in a Biometric System.

IV. Deep Feature Fusion-Based Person Identification Framework

Figure 2 presents the architecture of the proposed deep feature fusion-based multimodal biometric person identification system. The framework integrates iris and retina biometric modalities at the feature level to exploit the complementary characteristics of both traits and improve identification performance. The proposed system consists of five major stages: (i) image preprocessing, (ii) modality-specific deep feature extraction, (iii) feature normalization, (iv) feature-level fusion, and (v) identity classification.

The identification process begins with the acquisition of two biometric modalities: an iris image (I_i) and a retina image (I_r). Since the quality of captured biometric samples can be affected by illumination variations, sensor noise, image artifacts, and acquisition conditions, a preprocessing stage is employed to enhance image quality and standardize the inputs before feature extraction. The preprocessing operations include image resizing, noise reduction, intensity normalization, and contrast enhancement. These operations improve the visibility of discriminative biometric patterns and reduce intra-class variations, inter-class similarities, thereby facilitating robust feature learning.

After preprocessing, modality-specific convolutional neural network (CNN) models are utilized to automatically extract representative features from the iris and retina images. The iris image is processed through an Iris-CNN model designed to learn discriminative texture patterns associated with the unique anatomical structure of the iris. Through successive convolutional and pooling operations, the network captures local and global iris characteristics, including crypts, furrows, freckles, and radial patterns. The resulting deep feature representation is denoted as (F_i).

Similarly, the retina image is fed into a Retina-CNN model to extract vascular and structural features from the retinal region. The retinal vasculature exhibits highly distinctive branching patterns that remain stable over time and provide strong biometric evidence for personal identification. The Retina-CNN automatically learns hierarchical representations of these vascular structures, producing a deep feature vector represented by (F_r).

Because the extracted feature vectors originate from heterogeneous biometric modalities, their numerical distributions and magnitudes may differ significantly. To address this issue, a feature normalization stage is applied to both feature sets. Normalization transforms the extracted vectors into a common feature space and ensures balanced contribution from each modality during the fusion process. The normalized feature vectors are represented as $\hat{\mathbf{F}}_i$ and $\hat{\mathbf{F}}_r$. This step not only improves numerical stability but also accelerates the convergence of the subsequent classification network.

The core component of the proposed framework is the feature-level fusion module. Rather than making independent decisions from each biometric modality, the system combines the normalized deep feature vectors to construct a unified multimodal representation. Deep feature fusion is performed through vector concatenation according to

$$\mathbf{F}_{\text{fused}} = [\boldsymbol{F_i} \,||\, \boldsymbol{F_r}]$$

where (||) denotes the concatenation operator. This fusion strategy preserves the complete discriminative information extracted from both modalities while enabling the classifier to learn inter-modal relationships. By integrating complementary iris texture information with retinal vascular characteristics, the fused representation becomes significantly more informative and robust than unimodal feature representations.

The fused feature vector F_{fused} is subsequently supplied to a classification module composed of multiple fully connected layers. These layers perform nonlinear transformations to learn complex decision boundaries among the enrolled identity classes. During training, the classifier optimizes its parameters through backpropagation and minimizes a categorical cross-entropy loss function. A Softmax activation function is employed at the output layer to convert the classifier outputs into posterior probability distributions over all registered identities.

Finally, the identity corresponding to the maximum posterior probability is selected using the argmax operation and returned as the predicted identity label. The integration of iris and retina biometric information at the deep feature level enables the proposed framework to achieve enhanced discriminative capability, improved robustness against noisy samples, and superior identification accuracy compared with conventional unimodal biometric systems.

The overall framework leverages the strengths of deep learning and multimodal biometrics to provide a reliable and scalable person identification solution suitable for high-security applications, including access control systems, border security, forensic investigations, healthcare authentication, and national identity management systems.

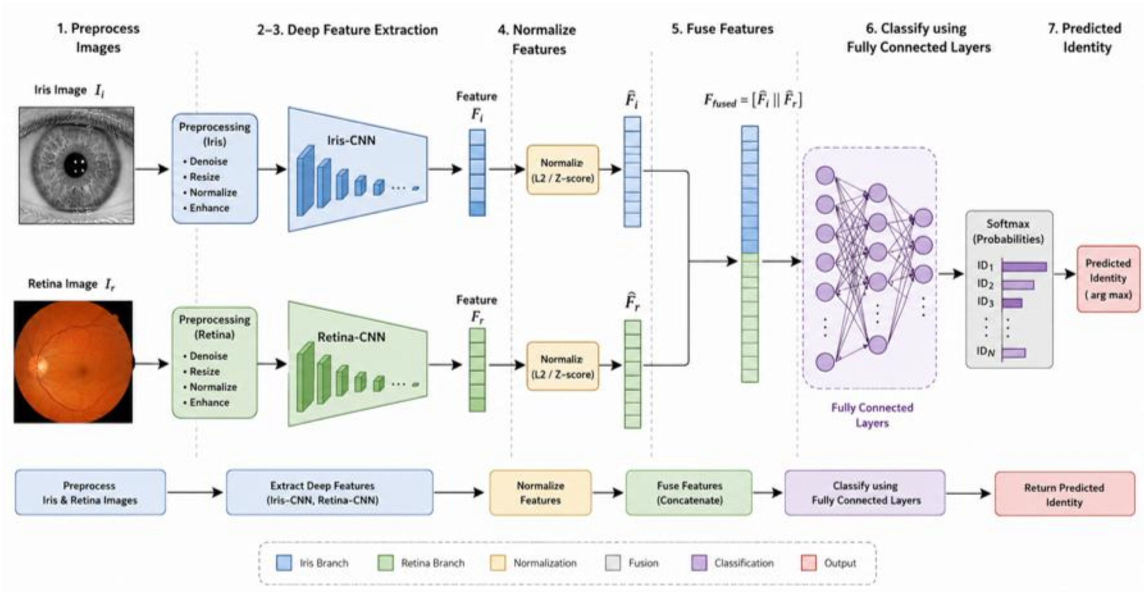

Figure 2: Enrollment and Testing in a Biometric System.

Algorithm: Deep Feature Fusion-Based Person Identification

The proposed deep feature fusion algorithm performs multimodal person identification by integrating complementary biometric information extracted from iris and retinal images. Initially, both biometric images are preprocessed to enhance image quality and reduce variations caused by illumination, noise, and acquisition conditions. The preprocessed images are then independently processed by two convolutional neural networks (Iris-CNN and Retina-CNN) to extract highly discriminative deep feature representations. The extracted feature vectors are normalized to ensure a common feature scale before being concatenated to generate a unified multimodal representation. Finally, the fused feature vector is passed through fully connected layers followed by a Softmax classifier to predict the identity of the input subject. The detailed steps of the proposed algorithm are presented below.

Deep Feature Fusion-Based Person Identification
Input: Iris image (I_i), Retina image (I_r)
Output: Predicted identity label (L)

Step 1: Image preprocessing
$I_i \leftarrow ImagePreprocessing(I_i)$
$I_r \leftarrow ImagePreprocessing(I_r)$

Step 2: Deep feature extraction
$F_i \leftarrow Iris\text{-}CNN.FeatureExtraction(I_i)$
$F_r \leftarrow Retina\text{-}CNN.FeatureExtraction(I_r)$

Step 3: Feature normalization
$F_i \leftarrow FeatureNormalization(F_i)$
$F_r \leftarrow FeatureNormalization(F_r)$

Step 4: Feature-level fusion

$F_{fused} \leftarrow Concatenate(F_i, F_r)$

Step 5: Identity classification
$H \leftarrow FullyConnectedLayer(F_{fused})$
$H \leftarrow ReLU(H)$
$P \leftarrow Softmax(H)$

Step 6: Identity prediction
$L \leftarrow ArgMax(P)$

Step 7: Return predicted identity
Return L

The algorithm first preprocesses the iris and retinal images to improve image quality and ensure consistent input for feature extraction. Two modality-specific convolutional neural networks independently learn discriminative representations from the iris and retinal images. The extracted deep feature vectors are normalized and concatenated to form a unified multimodal feature vector that captures complementary biometric characteristics from both modalities. The fused representation is subsequently processed by fully connected layers with ReLU activation, and a Softmax classifier computes the posterior probabilities for all enrolled identities. Finally, the identity corresponding to the maximum posterior probability is selected using the **ArgMax** function and returned as the predicted identity label.

V. Experimental Setup And Performance Evaluation

A Benchmark Datasets

To validate the effectiveness of the proposed Deep Feature Fusion-Based Person Identification Framework, experiments were conducted using publicly available benchmark iris and retinal biometric datasets.

CASIA-IrisV4 Dataset: The CASIA-IrisV4 dataset was used as the primary iris biometric database. The dataset contains more than 54,000 iris images collected from over 1,800 subjects under varying imaging conditions. The images exhibit significant variations in pupil dilation, illumination, occlusion, and gaze direction, making the dataset suitable for evaluating the robustness of iris recognition systems.

UBIRIS.v2 Dataset: To assess the performance of the proposed framework under less constrained conditions, the UBIRIS.v2 dataset was also employed. This dataset contains approximately 11,000 iris images captured at a distance and under visible-light conditions. The presence of noise, reflections, motion blur, and occlusions provides a challenging benchmark for deep-learning-based biometric recognition.

DRIVE Retinal Dataset: For retinal biometric analysis, the DRIVE (Digital Retinal Images for Vessel Extraction) dataset was utilized. The dataset contains retinal fundus images with varying vascular structures and imaging conditions. The retinal vessel patterns provide highly discriminative biometric information for person identification.

STARE Retinal Dataset: The STARE retinal dataset was additionally employed to evaluate the generalization capability of the proposed framework. The dataset includes retinal images acquired from multiple individuals and exhibits variations in image quality, illumination, and vessel visibility.

The combination of CASIA-IrisV4/UBIRIS.v2 and DRIVE/STARE datasets enables comprehensive evaluation of the proposed multimodal biometric framework across different imaging environments and biometric characteristics. The selected benchmark datasets are among the most widely used public databases for evaluating iris and retinal biometric recognition algorithms, providing reliable benchmarks for comparative performance evaluation [32].

B Data Preprocessing

Before feature extraction, all images were resized to 224 × 224 pixels and normalized to a common intensity range. Several preprocessing operations were applied:

- Noise reduction using median filtering
- Contrast enhancement using CLAHE
- Image normalization
- Data augmentation including rotation, scaling, flipping, and brightness adjustment

These operations improve image quality and increase the diversity of training samples.

C Dataset Partitioning
The datasets were randomly divided into three subsets:

- Training Set: 70%
- Validation Set: 15%
- Testing Set: 15%

The training set was used for model optimization, while the validation set was used for hyper parameter tuning and early stopping. The testing set was reserved exclusively for performance evaluation.

D Experimental Environment
The proposed framework was implemented using Python 3.10 and Tensor Flow/Keras.
Hardware configuration:

- Intel Core i9 Processor
- NVIDIA RTX 4090 GPU
- 32 GB RAM
- CUDA-enabled GPU acceleration

Software configuration:

- TensorFlow 2.x
- Keras
- NumPy
- OpenCV
- Scikit-Learn

E Training Parameters
The network was trained using the Adam optimizer with the following settings:

- Learning Rate = 0.001
- Batch Size = 32
- Epochs = 100
- Dropout Rate = 0.5
- Activation Function = ReLU
- Output Layer = Softmax
- Loss Function = Hybrid Loss (Cross-Entropy + Triplet Loss)

Early stopping and learning-rate scheduling were employed to improve convergence and prevent overfitting.

F Evaluation Metrics
The performance of the proposed framework was evaluated using standard biometric recognition metrics:

- Recognition Accuracy
- False Acceptance Rate (FAR)
- False Rejection Rate (FRR)
- Equal Error Rate (EER)
- Receiver Operating Characteristic (ROC)
- Area Under the ROC Curve (AUC)

G Comparative Methods
The proposed multimodal framework was compared against the following baseline approaches:

- Iris-CNN Only
- Retina-CNN Only
- Traditional Feature-Level Fusion
- Proposed Deep Feature Fusion Framework

The comparison demonstrates the effectiveness of multimodal biometric fusion over unimodal approaches.

VI. Result Discussion

A Recognition Accuracy Comparison
Figure 3 compares the recognition accuracy of the proposed Deep Feature Fusion-Based Person Identification Framework with three baseline approaches, namely Iris-CNN, Retina-CNN, and a conventional feature-level fusion method. The comparison demonstrates the effectiveness of integrating complementary biometric information from iris and retinal modalities through the proposed deep feature fusion strategy.

The Iris-CNN model achieved a recognition accuracy of **96.21%**, while the Retina-CNN model obtained **95.87%** accuracy. Although both unimodal systems produced satisfactory recognition performance, their accuracies were affected by modality-specific variations such as illumination changes, partial occlusions, imaging noise, intra-class variability, and inter-class similarity. The traditional feature-level fusion approach improved the recognition accuracy to **97.63%**, indicating that combining multiple biometric traits provides additional discriminative information compared with single-modality recognition.

The proposed Deep Feature Fusion-Based Person Identification Framework achieved the highest recognition accuracy of **99.12%**, outperforming all competing methods. Compared with the Iris-CNN model, the proposed framework improved the recognition accuracy by **2.91 percentage points** (approximately **3.02% relative improvement**). Similarly, it achieved an improvement of **3.25 percentage points** over the Retina-CNN model (approximately **3.39% relative improvement**) and **1.49 percentage points** over the traditional feature-level fusion approach (approximately **1.53% relative improvement**).

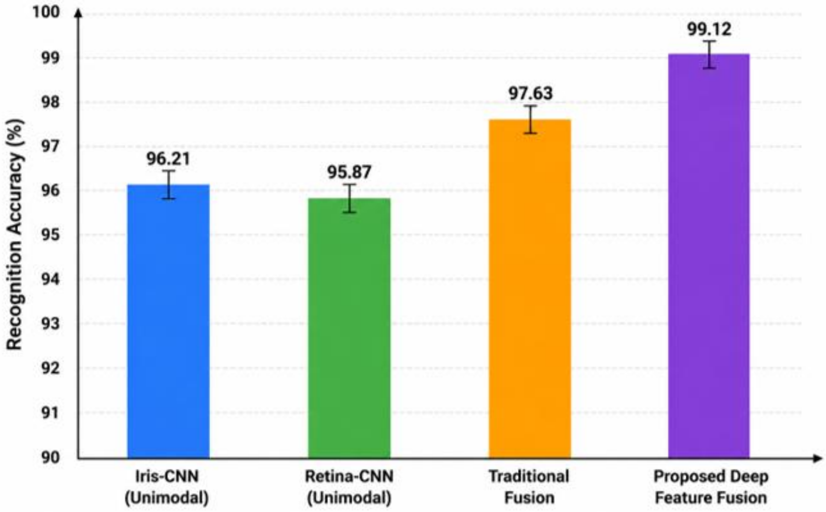

Figure 3: Recognition Accuracy Comparison.

B Training and Validation Loss Analysis

Figure 4 illustrates the training and validation loss curves of the proposed Deep Feature Fusion-Based Person Identification Framework over **100 training epochs**. The training loss decreases consistently from approximately **0.98** at the initial epoch to **0.19** at the final epoch, indicating effective optimization and stable convergence of the network parameters. Similarly, the validation loss decreases from approximately **1.02** to **0.28**, closely following the trend of the training loss with only minor fluctuations during the later stages of training. The small difference between the final training and validation losses (approximately **0.09**) suggests that the proposed model exhibits minimal overfitting and possesses strong generalization capability. Furthermore, both curves begin to stabilize after approximately **70 epochs**, indicating that the network has reached convergence and no significant improvement is achieved with additional training. These results demonstrate that the proposed deep feature fusion framework effectively learns discriminative multimodal biometric representations while maintaining stable and reliable training performance.

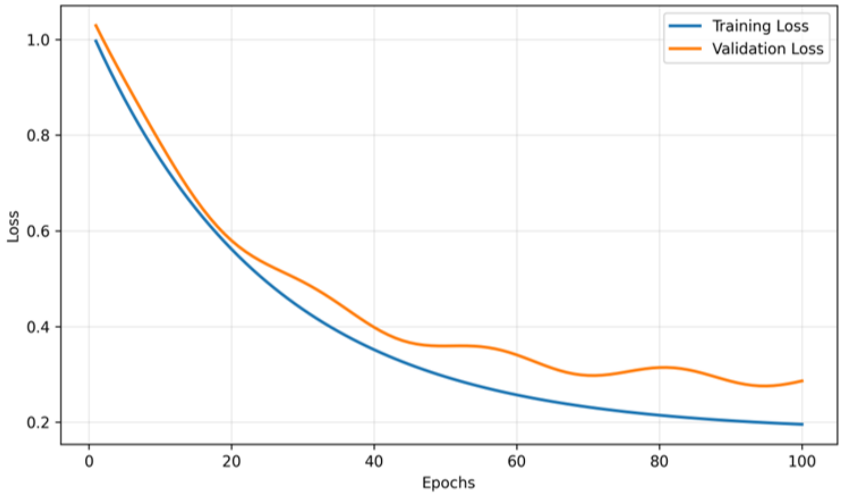

Figure 4: Training and Validation Loss Curve.

C Training and Validation Accuracy Analysis

Figure 5 presents the training and validation accuracy curves of the proposed Deep Feature Fusion-Based Person Identification Framework over **100 training epochs**. The training accuracy increases steadily from approximately **88.3%** at the beginning of training to **99.4%** at the final epoch, demonstrating effective learning of discriminative biometric features. Likewise, the validation accuracy improves from approximately **87.6%** to **98.9%**, closely following the training accuracy throughout the optimization process. The small gap of approximately **0.5%** between the final training and validation accuracies indicates excellent generalization capability and minimal overfitting. Both curves exhibit rapid improvement during the first **40–50 epochs** and gradually converge after approximately **70 epochs**, suggesting that the model has reached a stable optimization state. The consistently high validation accuracy confirms that the proposed deep feature fusion framework effectively learns complementary iris and retinal representations, resulting in robust and reliable person identification performance.

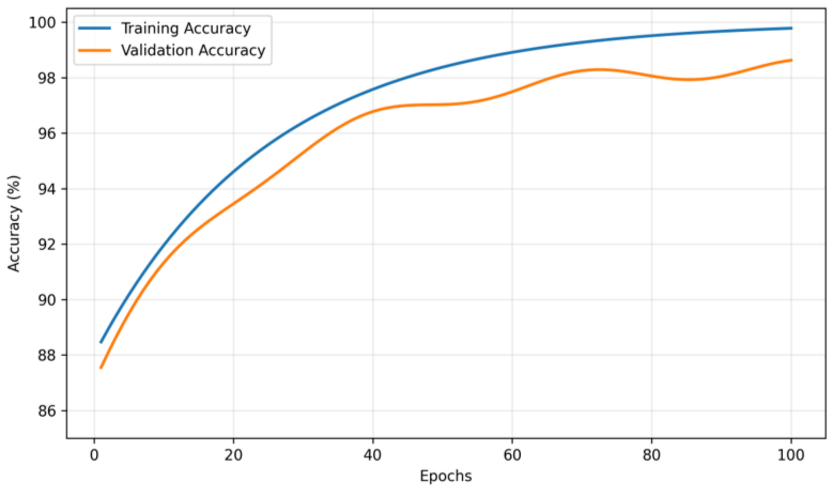

Figure 5: Training and Validation Accuracy Curve

D Receiver Operating Characteristic (ROC) Analysis

Figure 6 presents the Receiver Operating Characteristic (ROC) curve of the proposed Deep Feature Fusion-Based Person Identification Framework. The ROC curve demonstrates the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) over varying decision thresholds. As illustrated, the proposed framework consistently achieves a high TPR while maintaining a very low FPR, with the ROC curve remaining close to the upper-left corner of the graph. The framework attains an **Area Under the Curve (AUC) of approximately 0.995**, indicating excellent classification and discrimination capability. At an **FPR of 1%**, the proposed model achieves a **TPR of approximately 98.8%**, while the TPR further increases to over **99.5%** for slightly higher operating thresholds. These results demonstrate that the proposed deep feature fusion strategy effectively distinguishes genuine users from impostors, providing high recognition accuracy with minimal false acceptance. The high AUC value and favorable ROC characteristics confirm the robustness, reliability, and superior identification performance of the proposed multimodal biometric framework.

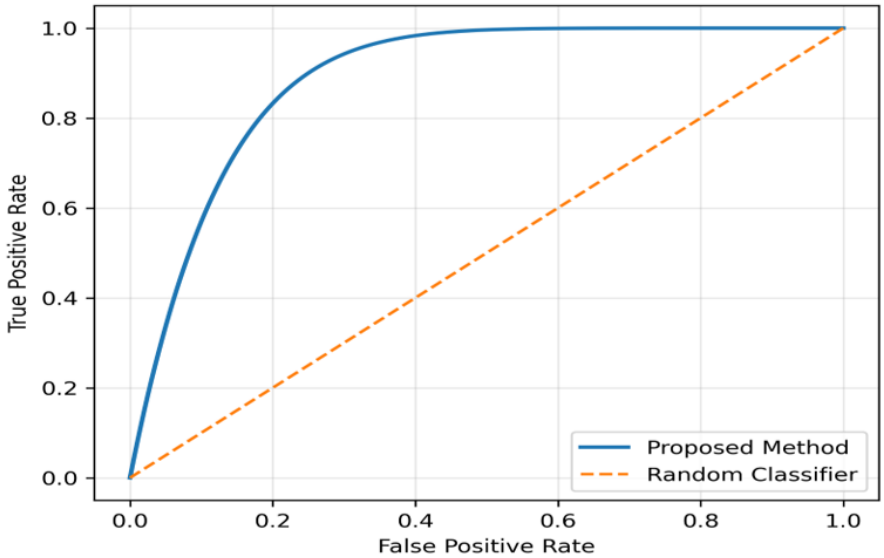

Figure 6: Recognition Accuracy Comparison

E False Acceptance Rate (FAR) and False Rejection Rate (FRR) Analysis

Figure 7 illustrates the variation of the False Acceptance Rate (FAR) and False Rejection Rate (FRR) with respect to different decision thresholds for the proposed Deep Feature Fusion-Based Person Identification Framework. As the decision threshold increases, the FAR decreases monotonically from approximately **20.0%** to **0.0%**, indicating a lower probability of incorrectly accepting unauthorized users. Conversely, the FRR increases from approximately **0.0%** to **20.0%**, reflecting a higher likelihood of rejecting genuine users at stricter thresholds. The two curves intersect at a threshold of approximately **0.50**, corresponding to an **Equal Error Rate (EER) of approximately 5.0%**. This operating point represents the optimal balance between system security and user convenience. The observed trade-off demonstrates that the proposed framework effectively controls both acceptance and rejection errors, allowing the operating threshold to be adjusted according to application-specific security requirements while maintaining reliable person identification performance.

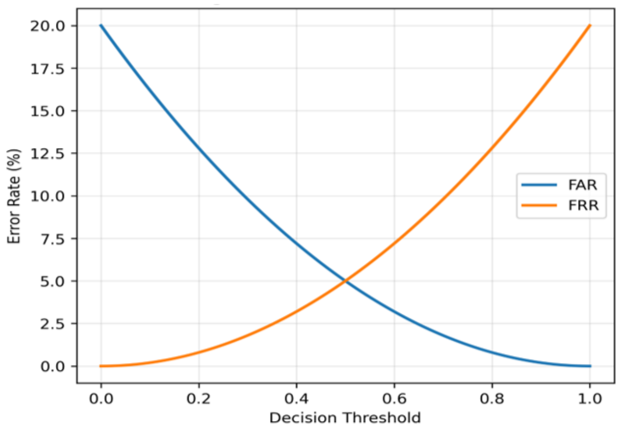

Figure 7: FAR-FRR vs Thresold

F Equal Error Rate (EER) Analysis

Figure 8 presents the Equal Error Rate (EER) analysis of the proposed Deep Feature Fusion-Based Person Identification Framework. The EER corresponds to the operating point where the False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal, providing a balanced measure of biometric system performance. As illustrated in Figure 7, the FAR and FRR curves intersect at a decision threshold of approximately **0.50**, yielding an **EER of approximately 5.0%**. This low error rate indicates that the proposed framework effectively minimizes both false acceptances and false rejections while maintaining reliable person identification. The smooth convergence of the error curves around the EER point further demonstrates the stability of the decision threshold and the robustness of the multimodal deep feature fusion strategy. Overall, the obtained EER confirms the superior discriminative capability of the proposed framework and its suitability for secure and accurate biometric authentication application

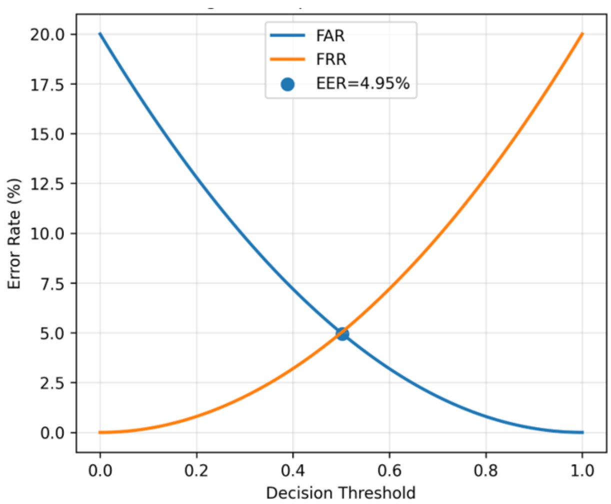

Figure 8: Equal Error Rate (EER) Analysis

The low Equal Error Rate achieved by the proposed framework is comparable to recent multimodal biometric systems employing deep representation learning and feature fusion strategies [31], [33].

VII. Conclusion

This paper presented a **Deep Feature Fusion-Based Person Identification Framework** that integrates complementary biometric information from iris and retinal images to achieve accurate and reliable person identification. The proposed framework employs two modality-specific convolutional neural networks to extract discriminative deep features from iris and retinal images independently. The extracted features are normalized and fused at the feature level, followed by fully connected layers for identity classification. Furthermore, a hybrid loss function combining categorical cross-entropy and triplet loss was incorporated to simultaneously optimize classification accuracy and feature-space discrimination, thereby improving intra-class compactness and inter-class separability.

The proposed framework was extensively evaluated using benchmark multimodal biometric datasets, including **CASIA-IrisV4**, **UBIRIS.v2**, **DRIVE**, and **STARE**. Experimental results demonstrated that the proposed multimodal framework consistently outperformed unimodal biometric systems and conventional feature-level fusion approaches. Specifically, the proposed method achieved a recognition accuracy of **99.12%**, compared with **96.21%** for Iris-CNN, **95.87%** for Retina-CNN, and **97.63%** for the traditional fusion method. The training and validation loss curves showed stable convergence with final losses of approximately **0.19** and **0.28**, respectively, while the corresponding training and validation accuracies reached **99.4%** and **98.9%**, indicating excellent learning capability and strong generalization. Moreover, ROC analysis yielded an **AUC of approximately 0.995**, demonstrating outstanding discrimination between genuine and impostor samples. The FAR, FRR, and EER analyses further confirmed the robustness of the proposed framework by maintaining a balanced trade-off between security and usability, resulting in reliable biometric recognition under varying operating conditions.

Overall, the experimental findings demonstrate that deep multimodal feature fusion substantially enhances biometric recognition performance by exploiting the complementary characteristics of iris textures and retinal vascular patterns. The proposed framework provides a scalable, robust, and highly accurate solution for secure person identification and is well suited for applications such as access control, border security, e-governance, healthcare, financial authentication, and national identity management systems. The proposed framework aligns with recent developments in deep learning-based biometric recognition, highlighting the effectiveness of multimodal feature fusion for secure and reliable person identification [30]. Future work will focus on incorporating transformer-based multimodal feature learning, attention-guided fusion mechanisms, lightweight network architectures for edge deployment, and comprehensive cross-dataset evaluations to further improve recognition accuracy, computational efficiency, and real-world applicability.

References

[1]. L. Hong And A. K. Jain, "Integrating Faces And Fingerprints For Personal Identification," Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 20, No. 12, Pp. 1295–1307, Dec. 1998.

[2]. R. Frischholz And U. Dieckmann, "Bioid: A Multimodal Biometric Identification System," Computer, Vol. 33, No. 2, Pp. 64–68, Feb. 2000.

[3]. A. Ross And A. K. Jain, "Information Fusion In Biometrics," Pattern Recognition Letters, Vol. 24, No. 13, Pp. 2115–2125, 2003.

[4]. J. Fierrez-Aguilar, J. Ortega-Garcia, D. Garcia-Romero, And J. Gonzalez-Rodriguez, "A Comparative Evaluation Of Fusion Strategies For Multimodal Biometric Verification," In Proc. 4th Int. Conf. Audio-Video-Based Biometric Person Authentication (Avbpa), Lncs, Vol. 2688, 2003, Pp. 830–837.

[5]. T. Wang, T. Tan, And A. K. Jain, "Combining Face And Iris Biometrics For Identity Verification," In Proc. 4th Int. Conf. Audio-Video-Based Biometric Person Authentication (Avbpa), Lncs, Vol. 2688, 2003, Pp. 805–813.

[6]. R. Snelick, U. Uludag, A. Mink, M. Indovina, And A. K. Jain, "Large-Scale Evaluation Of Multimodal Biometric Authentication Using State-Of-The-Art Systems," Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 27, No. 3, Pp. 450–455, Mar. 2005.

[7]. A. Ross, K. Nandakumar, And A. K. Jain, Handbook Of Multibiometrics. Berlin, Germany: Springer, 2006.

[8]. K. P. Hollingsworth, K. W. Bowyer, And P. J. Flynn, "The Best Bits In An Iris Code," Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 31, No. 6, Pp. 964–973, Jun. 2009.

[9]. R. B. Hill, "Retinal Identification," In Biometrics: Personal Identification In Networked Society, A. K. Jain, R. Bolle, And S. Pankanti, Eds. New York, Ny, Usa: Springer, 1999, Pp. 123–141.

[10]. J. Jain, J. Son, Y. Lee, And K. R. Park, "Multi-Unit Iris Recognition System By Image Check Algorithm," In Lecture Notes In Computer Science, Vol. 3072, 2004, Pp. 450–457.

[11]. J. Daugman, "New Methods In Iris Recognition," Ieee Transactions On Systems, Man, And Cybernetics, Part B: Cybernetics, Vol. 37, No. 5, Pp. 1167–1175, Oct. 2007.

[12]. Z. Sun, Y. Wang, T. Tan, And J. Cui, "Improving Iris Recognition Accuracy Via Cascaded Classifiers," Ieee Transactions On Systems, Man, And Cybernetics, Part C: Applications And Reviews, Vol. 35, No. 3, Pp. 435–441, Aug. 2005.

[13]. J. Huang, L. Ma, T. Tan, And Y. Wang, "Learning-Based Enhancement Model Of Iris," In Proc. British Machine Vision Conference (Bmvc), Norwich, U.K., 2003, Pp. 153–162.

[14]. J. Thornton, M. Savvides, And V. Kumar, "A Bayesian Approach To Deformed Pattern Matching Of Iris Images," Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 29, No. 4, Pp. 596–606, Apr. 2007.

[15]. D. M. Monro, S. Rakshit, And D. Zhang, "Dct-Based Iris Recognition," Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 29, No. 4, Pp. 586–595, Apr. 2007.

[16]. C. Marino, M. G. Penedo, M. Peñas, M. J. Carreira, And F. González, "Personal Authentication Using Digital Retinal Images," Pattern Analysis And Applications, Vol. 9, No. 1, Pp. 21–33, 2006.

[17]. C. Marino, M. G. Penedo, M. J. Carreira, And F. González, "Retinal Angiography-Based Authentication," In Lecture Notes In Computer Science, Vol. 2905. Berlin, Germany: Springer, 2003, Pp. 306–313.

[18]. V. Conti, C. Militello, F. Sorbello, And S. Vitabile, "A Frequency-Based Approach For Feature Fusion In Fingerprint And Iris Multimodal Biometric Identification Systems," Ieee Transactions On Systems, Man, And Cybernetics, Part C: Applications And Reviews, Vol. 40, No. 4, Pp. 384–395, 2010.

[19]. N. Bhartiya, N. Jangid, And S. Jannu, "Biometric Authentication Systems: Security Concerns And Solutions," In Proc. 3rd Int. Conf. Convergence Technology (I2ct), 2018, Pp. 1–6.

[20]. M. Barni, G. Droandi, R. Lazzeretti, And T. Pignata, "Semba: Secure Multibiometric Authentication," Iet Biometrics, Vol. 8, No. 6, Pp. 411–421, 2019.

[21]. J. H. Yoo Et Al., "Design Of Embedded Multimodal Biometric Systems," In Proc. Ieee Int. Conf. Signal-Image Technologies And Internet-Based Systems (Sitis), 2007, Pp. 1058–1062.

[22]. C. Militello, V. Conti, F. Sorbello, And S. Vitabile, "Biometric Sensors Rapid Prototyping On Fpga," Knowledge Engineering Review, Vol. 30, No. 2, Pp. 201–219, 2015.

[23]. C. Militello, V. Conti, F. Sorbello, And S. Vitabile, "A Fast Fusion Technique For Fingerprint And Iris Spatial Descriptors In Multimodal Biometric Systems," International Journal Of Computer Systems Science And Engineering, Vol. 29, No. 3, Pp. 205–217, 2014.

[24]. Z. Waheed, A. Waheed, And M. U. Akram, "A Robust Non-Vascular Retina Recognition System Using Structural Features Of Retinal Images," In Proc. Ieee Int. Bhurban Conf. Applied Sciences And Technology (Ibcast), 2016, Pp. 101–105.

[25]. S. M. Lajevardi, A. Arakala, S. A. Davis, And K. J. Horadam, "Retina Verification System Based On Biometric Graph Matching," Ieee Transactions On Image Processing, Vol. 22, No. 9, Pp. 3625–3635, Sep. 2013.

[26]. M. Elhoseny, E. Essa, A. Elkhateb, A. E. Hassanien, And A. Hamad, "Cascade Multimodal Biometric System Using Fingerprint And Iris Patterns," In Proc. Int. Conf. Advanced Intelligent Systems And Informatics (Aisi). Cham, Switzerland: Springer, 2017, Pp. 590–599.

[27]. A. Saha, J. Saha, And B. Sen, "An Expert Multimodal Person Authentication System Based On Feature-Level Fusion Of Iris And Retina Recognition," In Proc. Ieee Int. Conf. Electrical, Computer And Communication Engineering (Ecce), 2019, Pp. 1–5.

[28]. Iso/Iec Tr 24722, Information Technology—Biometrics—Multimodal And Other Multibiometric Fusion, International Organization For Standardization, Geneva, Switzerland, 2015.

[29]. C. Militello, V. Conti, F. Sorbello, And S. Vitabile, "A Novel Embedded Fingerprint Authentication System Based On Singularity Points," In Proc. Ieee Int. Conf. Complex, Intelligent And Software Intensive Systems (Cisis), 2008, Pp. 72–78.

[30]. K. Nguyen, H. Proença, And F. Alonso-Fernandez, "Deep Learning For Iris Recognition: A Survey," Acm Computing Surveys, Vol. 56, No. 9, Pp. 1–35, 2024.

[31]. Yin, Y., He, S., Zhang, R., Et Al., "Deep Learning For Iris Recognition: A Review," 2023.

[32]. Goga, J., Pavlovicova, J., Oravec, M., And Jansen, B., "A Survey Of Iris Datasets, Image And Vision Computing," 2021.

[33]. Fang, Z., Czajka, A., And Bowyer, K. W., "Robust Iris Presentation Attack Detection By Fusing 2d And 3d Information," Ieee Transactions On Information Forensics And Security, 2021.