

Optimized Feature-Level Fusion Model For Multimodal Biometric System

Bharathi R.^{1,*} and M. B. Anandaraju^{2,*}

¹Research Scholar, Assistant Professor, ECE, BGS Institute of Technology, Adichunchanagiri University, Mandya, India, 571448

²Professor, Department of ECE, BGS Institute of Technology, Adichunchanagiri University, B G Nagara, India, 571448

*E-mail: bharathiraju23@gmail.com; Anandaraju-mb.anandaraju@yahoo.com

Received for publication
June 18, 2025

Abstract

Biometric recognition is the automatic identification of individuals based on their unique behavioral or biological traits, creating a direct link to their identity that cannot be lost, shared, or duplicated. A multimodal biometric system integrates multiple biometric traits to enhance identification accuracy and security. This research presents an optimum fusion model that incorporates iris, fingerprint, and ear biometrics, employing feature-level fusion methodology. Local binary patterns (LBP) have been used on individual biometric traits to extract unique texture features and the Bayesian optimized support vector machine feature fusion (BSFF) technique for efficient and accurate classification has been proposed. The suggested model improves the performance and robustness of biometric authentication, making it appropriate for high-security applications. Experimental studies show that the proposed method achieves an average equal error rate (EER) of 0.1478, NRMSE of 0.0983, and accuracy of 0.9343 using fivefold cross-validation.

Keywords

Bayesian optimization, biometric recognition, equal error rate, feature level fusion, hyperparameter, local binary patterns, multimodal

I. Introduction

Biometrics is a way to find out who someone is by looking at their unique traits. These traits, which are also known as “biometric traits,” can be broken down into two groups: physiological and behavioral [1]. Biometrics makes things safer, more comfortable, and harder to fake because they are harder to lose, copy, or fake [2]. The iris and finger vein are two of the most accurate hand-based biometrics for identifying people [3]. Because the finger vein system is hidden inside the finger, it is hard to fake [4]. In the same way, the eye has unique features that make it hard to copy and change. Based on how many biological traits are used, biometrics are either unimodal or multimodal [5] (Figure 1). Multimodal biometrics needs a way to combine or fuse the different biometric traits that are used.

For multimodal biometrics to work, the different biometric traits must be able to work together [6]. Multiple fusion methods are used in multimodal biometrics at the decision, score, feature, and picture levels. As the fusion level goes up, more information is added, but it becomes harder to combine [7]. Score-level fusion is the most popular method of fusion because it hits a good balance between making fusion easy and giving enough information to tell the difference between a True and an imposter case [8].

Most biometric systems are unimodal, which means they authenticate with only one source of data [9]. However, unimodal biometric systems have challenges such as noise in sensitive data, inter- and intra-class variations, a lack of universality, and susceptibility to spoofing attempts. To address these difficulties, multimodal biometrics can be utilized [10].

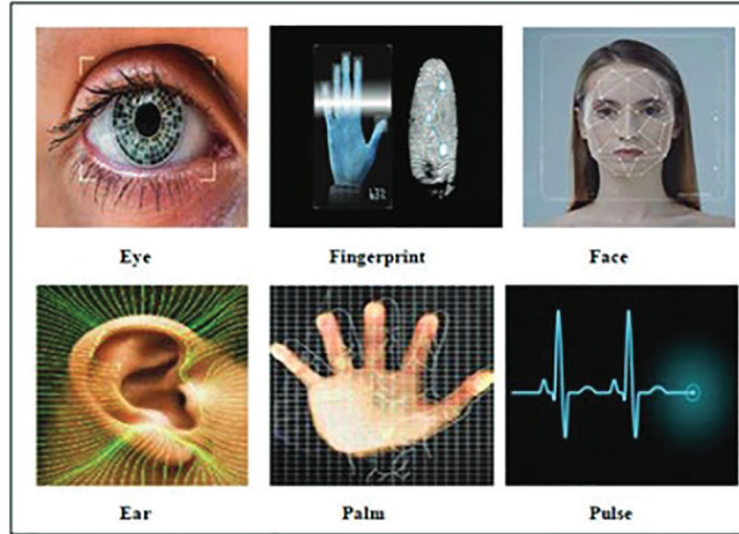


Figure 1: Various biometric traits.

Multimodal systems combine various biometric techniques, offering multiple sources of information for more secure authentication [11]. Furthermore, this study proposes an optimal fusion model that incorporates iris, fingerprint, and ear biometrics at the feature level. Local binary patterns (LBPs) are used for feature extraction, while concatenation-based feature-level fusion improves the system's discriminatory capabilities. The Bayesian Optimized SVM Feature Fusion (BSFF) technique is then used to ensure efficient and accurate classification. The study discusses existing multimodal recognition and fusion approaches in Section II, research contributions using Bayesian optimization (BO) in Section III, experimental validation in Section V, and finally, the results and future directions (Figure 1).

II. Literature Review

Multimodal biometric systems enhance authentication by integrating different traits [12]. This review focuses on multimodal biometric systems for various levels of fusion.

Chen et al. [13] improved data security and privacy by introducing a multimodal biometric recognition technique that uses federated learning. The method ensured decentralized data storage while maintaining a modest identification accuracy across modalities, such as voice, fingerprint, and face.

Kurdthongmee et al. [14] propose a fast and accurate pupil estimation method using a fine-tuned semantic segmentation model with a shallow convolutional neural network (CNN) backbone. It handles

varying lighting and camera conditions effectively, enabling real-time performance with high precision—ideal for applications in vision, biometrics, and medical diagnostics.

Kuklin et al. [15] address data drift in biometric systems using a Bayes–Minkowski-based feature correlation model and a neuroimmune approach. It employs adaptive learning to handle limited data and improves reliability, showing low equal error rate (EER) and strong resistance to adversarial attacks and information variability.

Adebayo et al. [16] state that Malay, as an under-resourced language, faces challenges in POS tagging due to limited annotated data. This study compares deep learning models and finds that a BERT + Bi-LSTM model achieves the highest performance, with 98.82% accuracy, outperforming traditional CRF and HMM approaches. The model excels in handling both known and unknown words, making it highly effective for POS tagging in Malay.

Bowyer et al. [17] presented an overview of 3D biometrics, stressing its superiority over 2D systems in handling occlusion and illumination challenges. They employed iterative closest point (ICP) algorithms for 3D face recognition, demonstrating improved recognition rates under varying environmental conditions.

Poh & Bengio [18] explored multimodal biometric fusion techniques, including voice, fingerprint, and iris integration. Using Bayesian fusion techniques, they achieved a 30% reduction in error rates compared with individual modalities.

Ross & Jain [19] provided a thorough assessment of multimodal biometrics, highlighting the integration

of complementary traits such as the face and iris. Their work utilized feature-level fusion with support vector machines (SVM), achieving better recognition accuracy.

Zhang et al. [20] studied palmprint-based biometric systems. They achieved a decent recognition rate on controlled datasets by using principal component analysis (PCA) and Gabor filters for feature extraction. They also highlighted the robustness of palmprint recognition when combined with other modalities.

Li & Jain examined the advantages of 3D biometrics over 2D systems, focusing on 3D face and ear recognition. Using geometric feature extraction algorithms, they demonstrated higher robustness against occlusion and age-related changes [21].

Chen et al. reviewed advancements in 3D ear biometrics, demonstrating its effectiveness in handling occlusion and providing stable features over time. They applied wavelet transform and 3D surface matching, on large-scale datasets [22].

Nguyen et al. [23] explored the integration of 3D biometrics in multimodal systems, combining face and ear recognition. By employing CNNs, they achieved an 88% improvement in robustness under varying environmental conditions.

While the primary focus of our work is on developing an accurate fusion model, several aspects of the proposed system support real-world deployment. First, texture-based LBP features are computationally lightweight and offer robustness to sensor inconsistencies, such as changes in resolution, lighting, and

imaging quality. Second, by combining three distinct biometric modalities, the system is inherently more secure against spoofing attempts, as attacking all modalities simultaneously is considerably more difficult. Finally, the use of widely accepted, non-contact biometrics (iris, fingerprint, and ear) makes the system more practical and user-friendly for high-security applications, where user cooperation and comfort are essential.

This research presents a novel multimodal biometric authentication system that uniquely combines iris, fingerprint, and ear traits using feature-level fusion, where LBP are applied uniformly for texture feature extraction and a Bayesian-optimized SVM fusion technique (BSFF) is introduced to enhance classification accuracy; the approach not only leverages the rich discriminative power of fused features but also ensures robustness, scalability, and superior performance.

III. Materials and Methods

The proposed feature-level fusion model for a multimodal biometric system combines iris, fingerprint, and ear data while extracting texture using LBP. The acquired features from each modality are concatenated to form a strong feature vector, which is then optimized using dimensionality reduction techniques, as illustrated in Figure 2. Finally, an SVM classifier and Bayesian Optimizer were utilized to achieve efficient and accurate recognition.

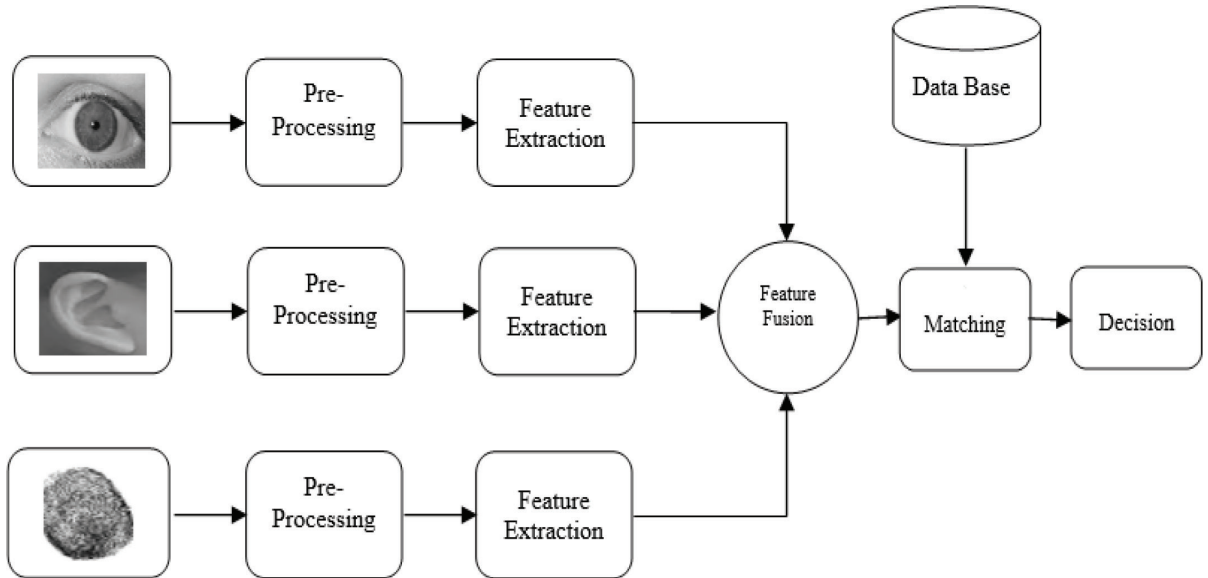


Figure 2: Multimodal Biometric System.

a. Dataset

Ear images from the IIT Delhi Ear Database v1.0, fingerprint images from CASIA-FingerprintV5, and iris images from CASIA with different resolutions were collected. The required dataset was prepared through preprocessing techniques, including resizing, data augmentation, normalization, and noise reduction, to ensure uniformity and enhance recognition performance.

b. LBPs

LBPs are employed for feature extraction from iris, ear, and fingerprint images due to their effectiveness in capturing texture details [24]. Iris-LBP extracts fine-grained texture patterns from iris images, enhancing feature robustness against illumination and scale variations [25]. Here, LBP captures the distinct structural and textural characteristics of the traits, improving its discriminative capability for biometric identification. Fingerprint-LBP encodes the patterns, preserving local texture information crucial for biometric recognition as mentioned in Figure 3.

The LBP operator encodes texture by comparing a pixel's intensity with its neighbors. Each neighbor is assigned a binary value, and 1 if its intensity is higher than or equal to the central pixel, otherwise 0 [26]. These values form a binary code, defining the LBP for the pixel, as expressed in Eqs (1) and (2).

$$\text{LBP}(S, t) = \sum_{j=0}^{n-1} I(s_j - s_m) 2^j \quad (1)$$

$$I(s) = \begin{cases} 1, & s \geq 0 \\ 0, & s < 0 \end{cases} \quad (2)$$

We acknowledge that feature-level concatenation can lead to high-dimensional vectors. To address this, we applied LBPs for texture feature extraction, which inherently produces compact and discriminative feature sets, while no additional dimensionality reduction (e.g., PCA) was applied in this study.

c. Fusion of fingerprint, ear, and iris

The extracted LBP features from all three modalities are then used for fusion and classification for improved recognition accuracy [27]. Feature-level fusion combines the extracted features for richer information and merges all the features for efficient decision-making [28]. Feature-level fusion by union was employed, combining the left (L) and right (R) features of the three traits and represented as in Eqs (3)–(5) for ear, iris, and fingerprint, respectively. Finally, these features are fused [29] into a single fused vector as in Eq. (6):

$$\text{FEj} = [\{\text{Re1}, \text{Re2}, \dots, \text{Ren}\} \{\text{Le1}, \text{Le2}, \dots, \text{Len}\}] \quad (3)$$

$$\text{FIj} = [\{\text{Ri1}, \text{Ri2}, \dots, \text{Rin}\} \{\text{Li1}, \text{Li2}, \dots, \text{Lin}\}] \quad (4)$$

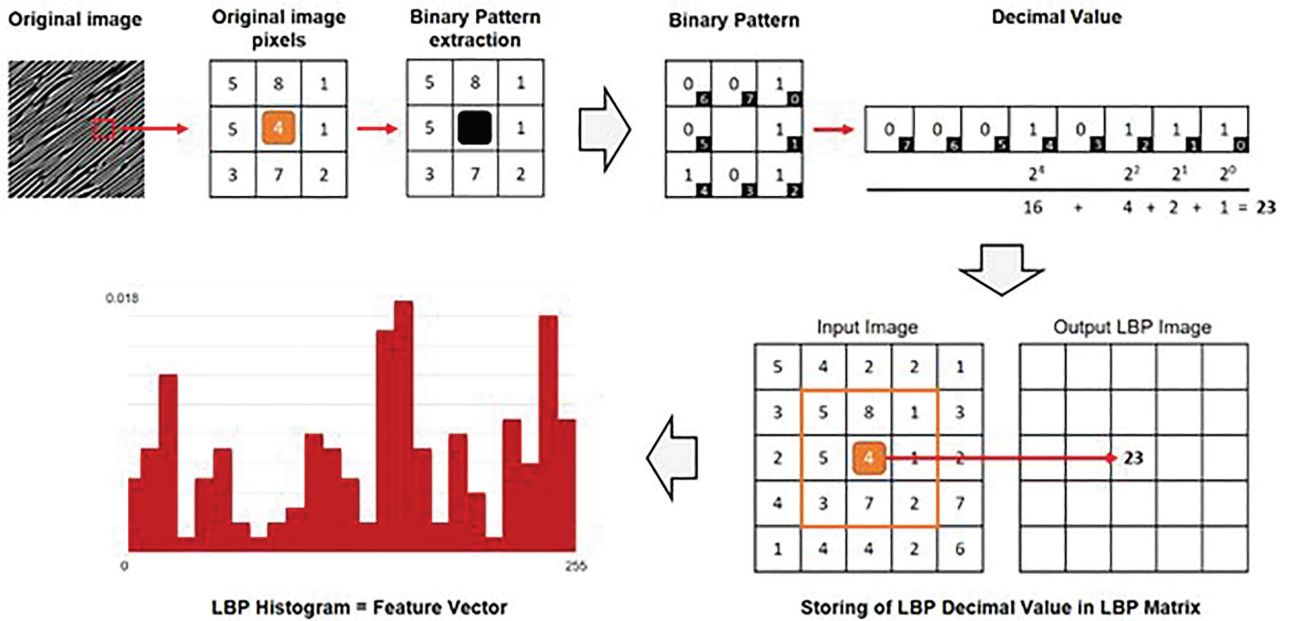


Figure 3: LBP. LBP, local binary pattern.

$$FFP_j = [Rfp1, Rfp2 \dots Rfp_n] \{Lfp1, Lfp2, \dots Lfp_n\} \quad (5)$$

$$F = [FE_j, FI_j, FFP_j] \quad (6)$$

d. Optimization and matching

The suggested work uses the data $\{(x_i, y_i)\}_{i=1}^N$ to train a machine learning model that learns or sets the parameters “ p ” using hyperparameters “ λ ” [30]. In essence, a machine learning problem aims to minimize [31], an objective function “ f ” using Eq. (7):

$$\text{minimize } f, p, \lambda \left(p, \lambda \mid \{(x_i, y_i)\}_{i=1}^N \right) \quad (7)$$

Eq. (8) is implemented to evaluate the proposed system with a linear kernel:

$$K_{\text{linear}}(x, x') = \langle x, x' \rangle \quad (8)$$

where (x, x') are the LBP vectors.

In summary, the hyperparameter “ λ ” of linear SVM is as in Eq. (9):

$$\lambda_{\text{LinearSVM}} = \{C\} \quad (9)$$

Furthermore, hyperparameter “ γ ” is optimized by Bayesian optimizer [32]. The optimized hyperparameters “ λ ” of linear SVM is given in Eq. (10):

$$\lambda_{\text{LinearSVM}} = \{C, \gamma\} \quad (10)$$

Where γ is the width of the kernel and C is the regularization hyperparameter [33]. BSFF multiplies optimized Bayesian weights with SVM prediction scores, sums them, and uses the result for match using Eq. (11).

$$F_s = \sum_{i=1}^n w_i \text{score}(x_i^i) \quad (11)$$

Where w denotes the score’s weight and F_s represents the fused score.

The score between the samples is evaluated from the Euclidean distance matcher using Eq. (12):

$$d(x, x') = \sqrt{\sum_{i=1}^D (x_i - x'_i)^2} \quad (12)$$

e. BO

BO is a type of sequential model-based optimization (SMBO) [34] that starts with sample hyperparameters to create a regression model [35]. An acquisition

function selects new samples, using the model [36] as a cost-effective proxy for the objective function in BSFF. Each sample x_i yields an observation of the objective function $y_i = f(x_i)$. The new sample is added to the set of all samples or data, $Q = \{(x_1, y_1) \dots (x_n, y_n)\}$, and is later utilized to update the model [37]. SMBO optimizes the LBP and SVM hyperparameters, to achieve better results and faster performance and is implemented in Algorithm 1. The BSFF algorithm is used to optimize the LBP hyperparameters, the SVM hyperparameters, and the weights of the feature level fusion using Algorithm 2.

Algorithm 1: SMBO

Data: f, x, A, M
 // Data initialisization
 $Q \leftarrow \text{Init. Samples } (f, x);$
 // run up to N steps
for $i \leftarrow |Q|$ to N **do**
 // Model Training
 $p(y \mid x, Q) \leftarrow \text{FitModel}(M, Q);$
 // Selection of best hyperparameter
 $x_i \leftarrow \text{argmax}_{x \in X} A(x, p(y \mid x, Q));$
 // estimate the hyperparameters
 $y_i \leftarrow f(x_i);$
 // append new data
 $Q \leftarrow Q \cup (x_i, y_i);$
end

Algorithm 2: BSFF Objective Function

Input: $F, y, \lambda_{\text{LBP}}, \lambda_{\text{svm}}$
Output: Statistical Parameter
 // hyperparameters usage from BO
 $\lambda_{\text{LBP}} \leftarrow \text{Trial from BO};$
 $\lambda_{\text{svm}} \leftarrow \text{Trial from BO};$
 // run for each Trait Ear, Iris, Fingerprint
for $j \leftarrow 1$ to n **do**
 // LBP feature extraction
 $F_{Ej} \leftarrow \text{Get LBP } (x_j, \lambda_{\text{LBP}});$ //Ear
 $FI_j \leftarrow \text{Get LBP } (x_j, \lambda_{\text{LBP}});$ // Iris
 $F_{FPj} \leftarrow \text{Get LBP } (x_j, \lambda_{\text{LBP}});$ // Fingerprint
 // concatenate the features as a Fused Feature vector
 $F \leftarrow \{F_{Ej}, F_{Ij}, F_{FPj}\};$
 // split Fused features (F) into train and test data

```

 $x_{train}, x_{test}, y_{train}, y_{test} \leftarrow \text{Train Test Split}(x, y);$ 
// train a SVM model with BO
SVMFit ( $x_{train}, y_{train}, \lambda_{SVM}$ );
// use the trained model (M) to predict test data
 $y_{pred} \leftarrow \text{SVM}_{\text{Predict}}(M, x_{test});$ 
// Evaluate Statistical Parameter (SP)
SP  $\leftarrow$  Get SP ( $y_{test}, y_{pred}$ );
return SP.

```

The BSFF approach innovates beyond traditional SVM-based multimodal fusion by performing early (feature-level) fusion of biometric traits, applying uniform LBP-based texture feature extraction, and incorporating a BO framework to automatically tune classifier parameters—resulting in a more robust, discriminative, and accurate biometric recognition system.

f. Performance evaluation

By utilizing feature-level fusion [38], an optimized multimodal biometric system was developed to improve resilience and accuracy by multimodal recognition [39]. Statistical parameters [40] are evaluated using the following established Eqs (13)–(18):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (13)$$

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

$$\text{F1 Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (16)$$

$$\text{EER} = 0.5 \times \{[\text{FP}/(\text{FP} + \text{TN})]/[\text{FN}/(\text{FN} + \text{TP})]\} \quad (17)$$

$$\text{NRMSE} = \sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N \frac{(y_i - y'_i)^2}{y_{\max} - y_{\min}}} \quad (18)$$

The actual and predicted values are represented by y_i and y'_i , respectively. The number of observations is denoted by N , and the maximum and minimum values of the actual data are $y_{\max}$ and $y_{\min}$.

IV. Results

In the research, computation has been done using a Intel Core i5-7th Generation CPU with 8GB

RAM. Datasets are collected from the Resource Collaboratory, and preprocessing has been carried out before the extraction of feature from all the three traits. In total, 3600 images were processed using data augmentation included mirroring, cropping, and resizing. Although the dataset includes around 3600 images, it is balanced across all modalities and classes, with each subject contributing iris, fingerprint, and ear samples. To address generalizability concerns, we used fivefold cross-validation, LBPs for compact feature extraction, and Bayesian-optimized SVM to reduce overfitting. Despite the dataset size, the model demonstrates strong performance and scalability potential for larger systems. Statistical parameters including true positive rate (TPR), false positive rate (FPR), and accuracy calculated for fivefold. Training and testing data were preprocessed and kept separately. Figure 4 depicts the sample image used in the proposed system. Table 1 presents the optimization process over 30 iterations, focusing on the objective function, runtime, best-obtained value, and hyperparameters like Box Constraint & Kernel Scale. The goal value increased from 0.096 to 0.069 between iteration 1 and iteration 7 and then stabilized, showing early convergence. Function evaluation times varied, indicating computational complexity owing to hyperparameter selection. Lower Box Constraints ≤ 0.001 and Kernel Scales ≤ 0.010 yielded the best results, emphasizing the importance of fine-tuning in this range for efficiency.

In the proposed work, BO algorithm performed 30 evaluations, yielding an observed objective function value of 0.06875 and an estimated value of 0.069326, indicating correct prediction (Table 2). Each evaluation lasted 12.3512 s. The Box Constraint 0.0013385 promotes flexibility and reduces misclassification penalties, whereas the Kernel Scale 0.0010021 produces highly localized decision boundaries and improves accuracy.

For effective multimodal analysis, the objective value should decline quickly at first, followed by steady convergence with minimum oscillations, as in Figure 5.

A parallel coordinates diagram is depicted in Figure 6, which illustrates the hyperparameters. The image offers a comprehensive view of the hyperparameters that produce optimal results, illustrating all optimization trials that the Bayesian optimizer has conducted across a variety of digits, datasets, and matchers. It is important to note that the majority of orientation and cell hyperparameter values are within the optimal range.

The effective selection of optimal parameters is facilitated by the convergence of hyperparameter optimization, thereby enhancing the performance of the model. In Figure 7, the hyperparameter optimization chart indicates a value of 0.0854 at Iteration 29, while the EER is 0.1478 as shown in Figure 8. This suggests a more refined equilibrium between the rates of false acceptance and false rejection. The improved EER serves as an illustration of the utility of feature-level fusion, which leads to a more precise and resilient multimodal biometric system.

Table 3 presents statistical analysis using fivefold validation, while Figure 9 illustrates the performance of the optimized multimodal system. The model's accuracy over five iterations is often high, with values ranging from 0.9101 to 0.9501 and an average of 0.9343. Iteration 4 achieves the highest accuracy of 0.9501, and Iteration 2 shows a 0.9101 a little drop. Similarly, the Macro AUC ranged from 0.8845 to 0.9644, indicating excellent classification performance, with Iteration 4 yielding the highest AUC. Macro Recall follows a similar pattern, peaking at 0.9559 in Iteration 4, demonstrating a strong sensitivity to good examples. Iteration 5 gets the greatest macro F1-score with 0.9570, showing balanced performance in terms of precision and recall. EER averages 0.1674, with Iteration 4 having the lowest EER of 0.1478, showing better classification threshold optimization. Finally, the NRMSE, which quantifies error, remains quite low across iterations, reaching a

peak of 0.0864 in Iteration 2, indicating greater model generalization.

Figures 10 and 11 depict the confusion matrix and ROC plot detailing the model's classification performance. The plotted curve of ROC lies consistently above the diagonal line, indicating the effectiveness of the proposed model in classifying different classes.

Table 4 presents a comparison between the proposed model and recent state-of-the-art approaches, highlighting a slight but meaningful improvement in accuracy over current methods.

Key challenges in real-world deployment include variations in acquisition conditions (e.g., lighting, sensor quality) and spoofing attempts. Our system mitigates these by using LBP, which are robust to illumination and texture changes, and by employing multimodal fusion (iris, fingerprint, ear), making it significantly harder for attackers to spoof all modalities simultaneously. This combination enhances resilience, security, and reliability in uncontrolled environments.

V. Conclusion

The proposed Optimized Feature Level Fusion Model enhances biometric recognition by integrating iris, fingerprint, and ear features using LBPs and Bayesian optimized SVM. The experimental results demonstrate its effectiveness in improving classification accuracy, making it a robust solution for multimodal

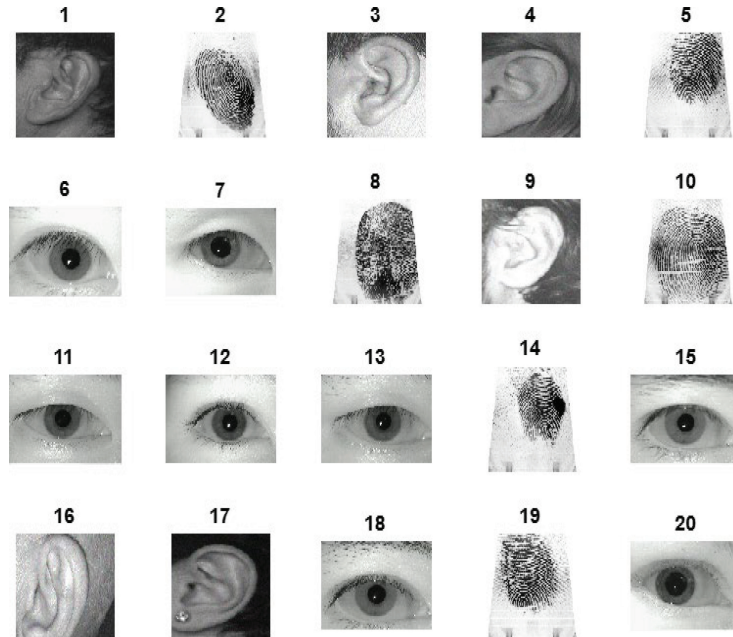


Figure 4: Sample Images used in Training process.

Table 1: Optimization analysis of the parameters

Iter	Eval	Objective	Objective (run time)	Best so far	Best so far est time	Box constraint	Kernel scale
1	B	0.096	13.135	0.096	0.096	0.337	0.002
2	B	0.079	1.576	0.079	0.082	871.77	0.010
3	A	0.331	91.062	0.079	0.092	431.78	0.029
4	A	0.833	2.125	0.079	0.079	0.002	26.961
5	A	0.079	3.016	0.079	0.079	0.074	0.002
6	A	0.902	0.735	0.079	0.079	0.041	998.72
7	B	0.069	12.351	0.069	0.069	0.001	0.001
8	A	0.077	2.731	0.069	0.069	0.065	0.004
9	A	0.223	46.893	0.069	0.069	152.81	0.001
10	A	0.167	0.636	0.069	0.079	0.001	0.004
11	A	0.079	5.504	0.069	0.078	964.37	0.001
12	A	0.079	2.760	0.069	0.069	984.19	0.003
13	A	0.079	2.050	0.069	0.069	7.319	0.004
14	A	0.098	20.025	0.069	0.069	0.059	0.001
15	A	0.763	1.779	0.069	0.069	0.001	0.039
16	A	0.894	242.51	0.069	0.069	8.479	984.37
17	A	0.079	2.163	0.069	0.068	0.455	0.005
18	A	0.398	0.693	0.069	0.069	0.001	0.748
19	A	0.079	7.152	0.069	0.068	90.169	0.001
20	A	0.098	6.362	0.069	0.069	0.001	0.002
21	A	0.079	2.122	0.069	0.069	984.74	0.007
22	A	0.196	55.945	0.069	0.069	863.67	0.005
23	A	0.073	5.081	0.069	0.069	0.001	0.001
24	A	0.831	195.47	0.069	0.069	965.59	8.748
25	A	0.094	1.953	0.069	0.069	0.001	0.002
26	A	0.079	4.070	0.069	0.069	840.91	0.002
27	A	0.079	7.255	0.069	0.069	0.062	0.001
28	A	0.079	2.759	0.069	0.069	0.195	0.003
29	A	0.240	0.550	0.069	0.069	0.001	0.135
30	A	0.388	149.79	0.069	0.069	920.34	0.236

Table 2: Hyper parameter optimization process using Bayesian approach

BO process	Hyperparameter tuning
Total function evaluations	30
Observed objective function value	0.06875
Estimated objective function value	0.069326
Function evaluation time	12.3512
Box constraint	0.0013385
Kernel scale	0.0010021

BO, Bayesian optimization.

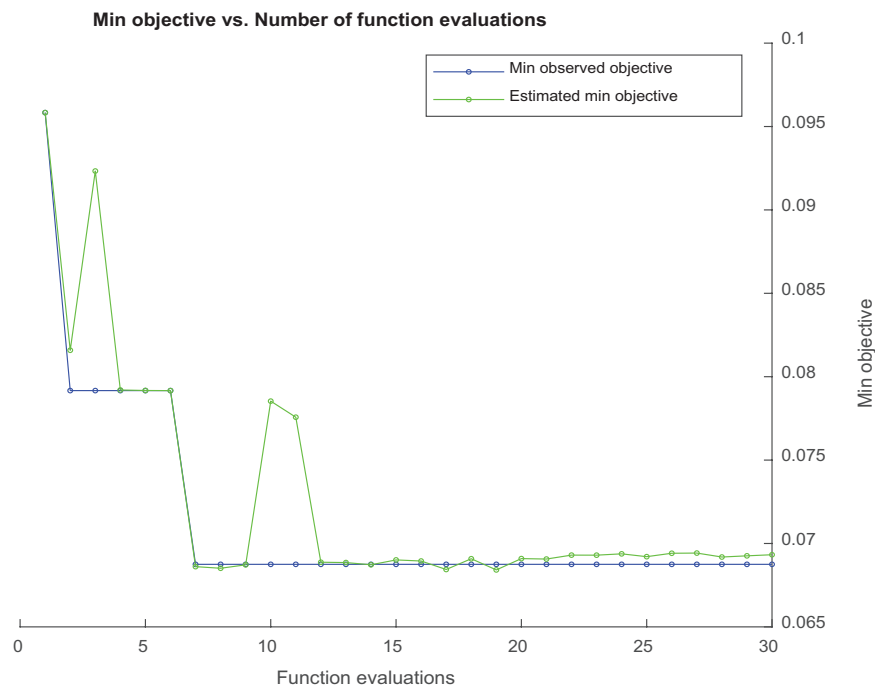


Figure 5: Acquisition function used in BO. BO, Bayesian optimization.

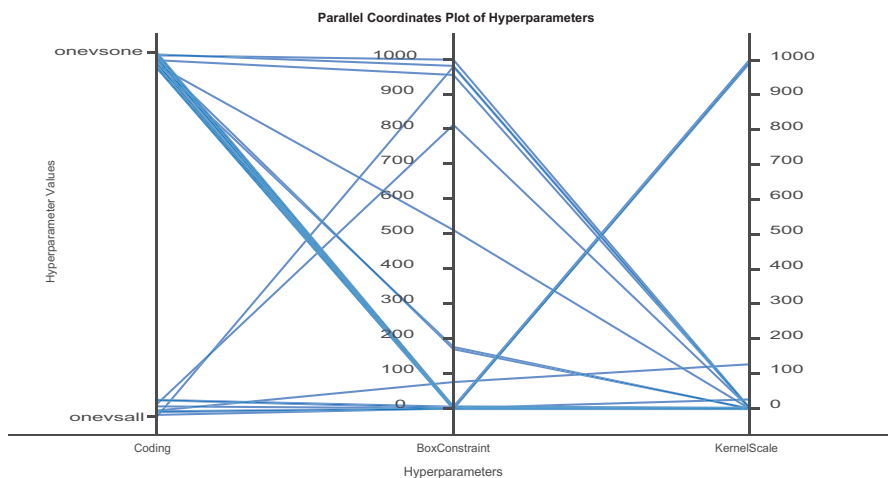


Figure 6: Parallel coordinates plot of the Hyperparameters.

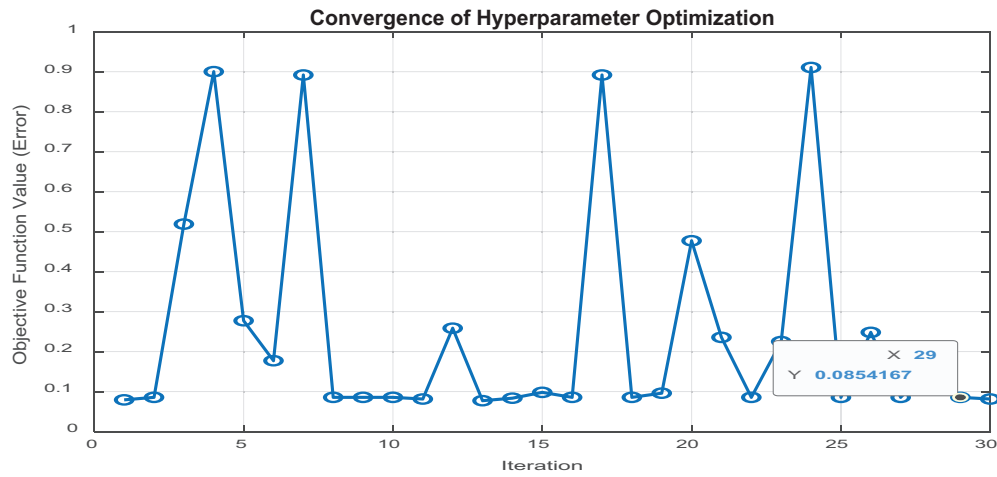


Figure 7: Convergence of hyper parameter optimization.

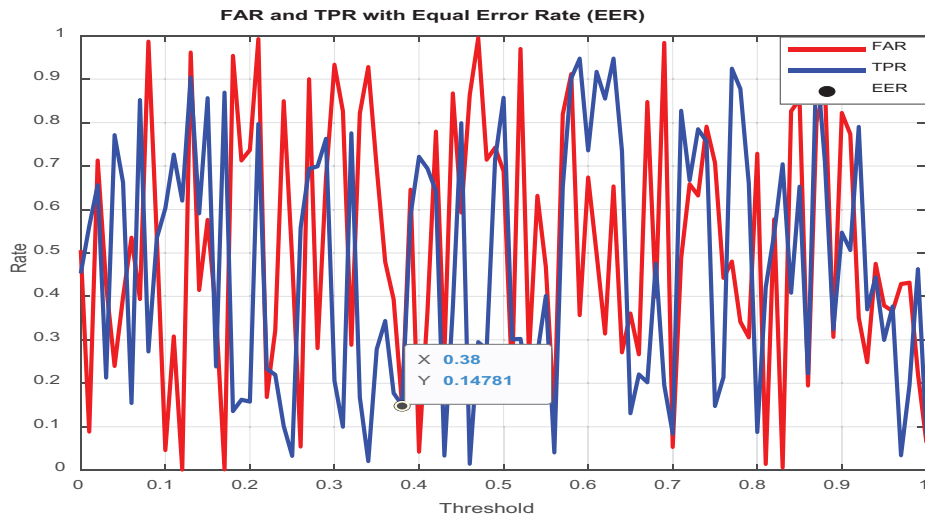


Figure 8: Optimized EER. EER, equal error rate; FPR, false positive rate; TPR, true positive rate.

Table 3: Statistical Parameter analysis using 5-Fold cross Validation.

Iteration	Accuracy	Macro AUC	Macro Recall	Macro F1-score	EER	NRMSE
1	0.9480	0.9133	0.8853	0.9131	0.1491	0.0897
2	0.9101	0.8946	0.8728	0.8937	0.2325	0.0864
3	0.9328	0.8845	0.9269	0.8740	0.1497	0.0879
4	0.9501	0.9644	0.9559	0.9494	0.1478	0.0983
5	0.9306	0.9161	0.9087	0.9570	0.1581	0.0941
Average	0.9343	0.9146	0.9099	0.9174	0.1674	0.0913

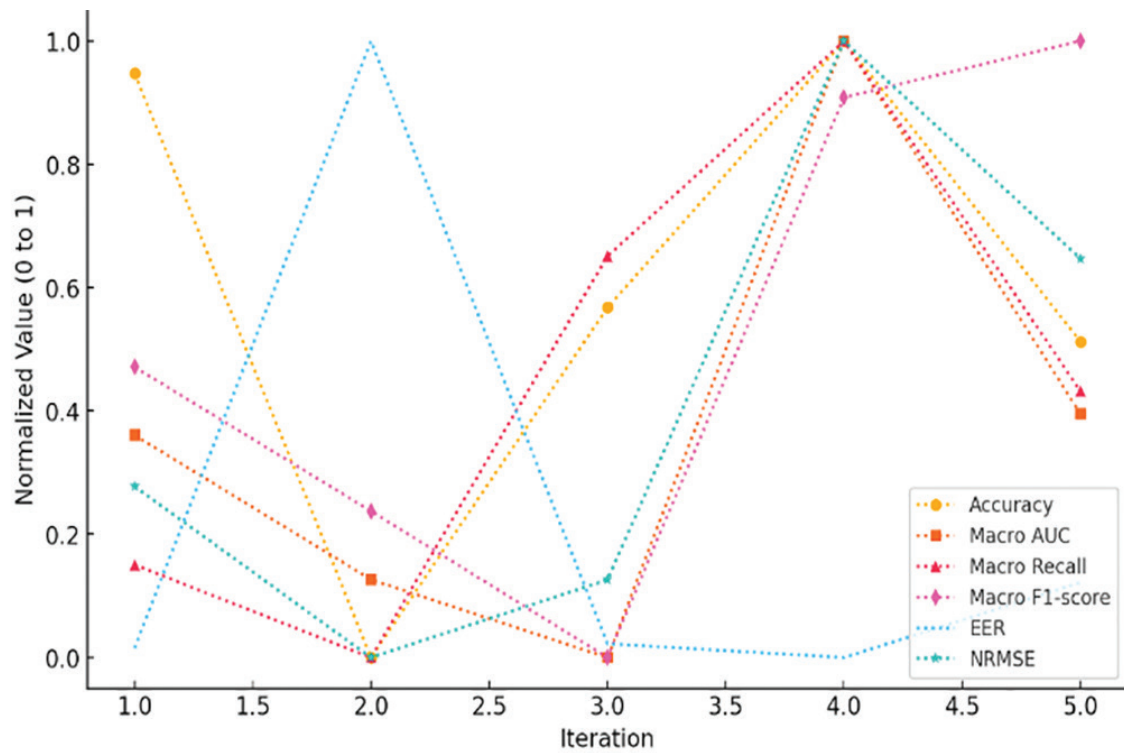


Figure 9: Performance of optimized multimodal system. EER, equal error rate.

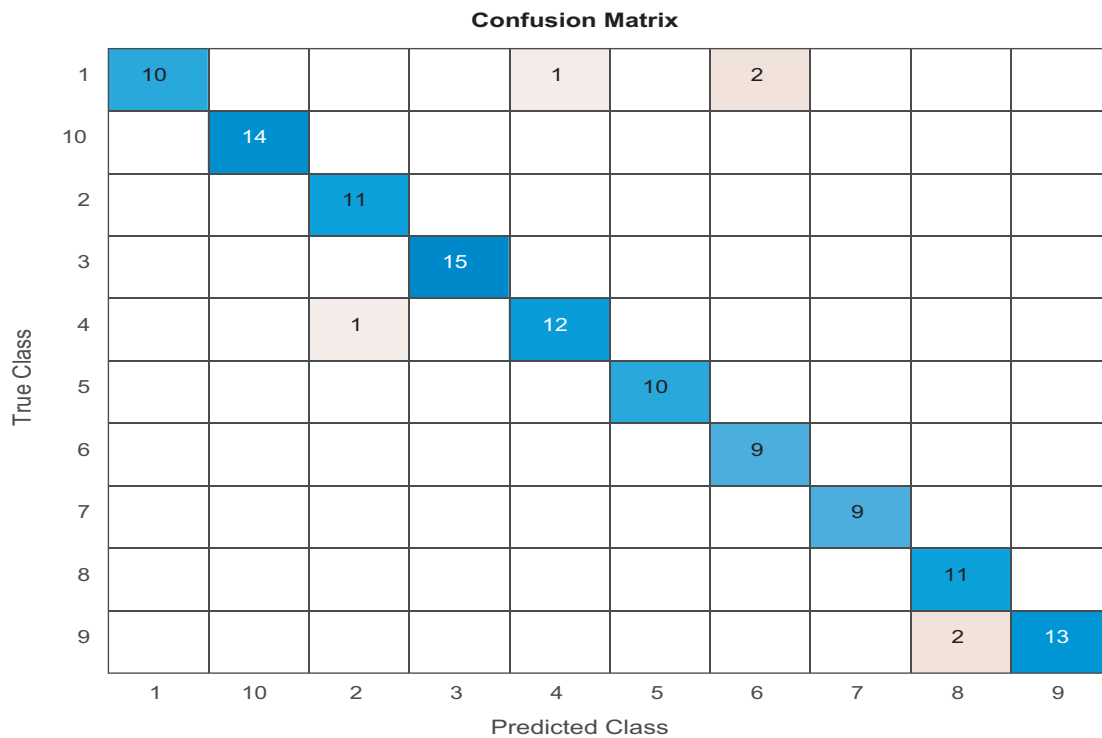


Figure 10. Confusion matrix in fourth iteration.

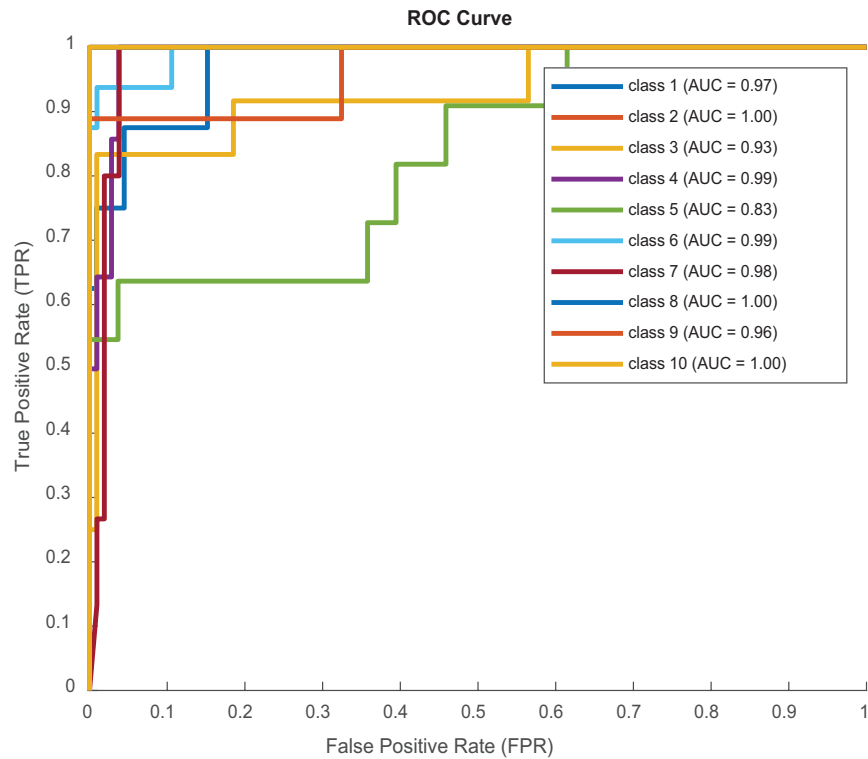


Figure 11. ROC curve in fourth iteration. FPR, false positive rate.

Table 4: Comparison of proposed model with state of art of models

No.	Authors	Accuracy in%
1	Kumarmohanta et al.[41]	85.20
2	Soleymani et al.[42]	88.12
3	Alshardan et al. [6]	90.01
4	Proposed Model	95.01

biometric systems. The objective function value of 0.06875 was obtained through BO, which also refined decision boundaries through an optimized Box Constraint of 0.0013 and Kernel Scale of 0.0010. This was achieved after 30 evaluations. Furthermore, Iteration 4 obtained the highest accuracy of 0.9501 and Macro AUC of 0.9644, thereby guaranteeing effective classification. Conversely, Iteration 2 reduced NRMSE errors. The model exhibited stability, with a balanced recall of 0.9559 and an F1-score of 0.9570. Additionally, the lowest EER of 0.1478 mitigated misclassification risks. These findings validate the model’s effectiveness and highlight its potential for biometric applications, including secure

authentication and identity verification. Future work will involve benchmarking our method against deep learning-based fusion approaches and conducting cross-dataset validation to assess the tradeoffs in accuracy, computational complexity, deployment feasibility, and the model’s generalizability across diverse real-world conditions.

References

- [1] Jasmine RM, Jasper J, Geetha MR. An efficient secure cryptosystem using improved identity-based encryption with multimodal biometric authentication and authorization in cloud environments. *Wireless Networks*. 2024 Jun 10; 1-21.
- [2] Amirthalingam G, Radhamani G. A multimodal approach for face and ear biometric system. *International Journal of Computer Science Issues (IJCSI)*. 2013 Sep 1;10(5):234.
- [3] Ma Y, Huang Z, Wang X, Huang K. An overview of multimodal biometrics using the face and ear. *Mathematical Problems in Engineering*. 2020;2020(1):6802905.

- [4] Byeon H, Raina V, Sandhu M, Shabaz M, Keshta I, Soni M, Matrouk K, Singh PP, Lakshmi TV. Artificial intelligence-enabled deep learning model for multimodal biometric fusion. *Multimedia Tools and Applications*. 2024 Oct;83(33): 80105-28.
- [5] Singh HK, Baranwal N, Singh KN, Singh AK. Using multimodal biometric fusion for watermarking of multiple images. *IEEE Transactions on Consumer Electronics*. 2024 Feb 29.
- [6] Alshardan A, Kumar A, Alghamdi M, Maashi M, Alahmari S, Alharbi AA, Almukadi W, Alzahrani Y. Multimodal biometric identification: leveraging convolutional neural network (CNN) architectures and fusion techniques with fingerprint and finger vein data. *PeerJ Computer Science*. 2024 Oct 31;10: e2440.
- [7] Huang Z, Liu Y, Li C, Yang M, Chen L. A robust face and ear based multimodal biometric system using sparse representation. *Pattern Recognition*. 2013 Aug 1;46(8):2156-68.
- [8] F. Ahamed, F. Farid, B. Suleiman, Z. Jan, L.A. Wahsheh and S. Shahrestani, An intelligent multimodal biometric authentication model for personalised healthcare services, *Future Internet*. 2022; 14(8); 222.
- [9] S.A. Abdulrahman and B. Alhayani et al., A comprehensive survey on the biometric systems based on physiological and behavioural characteristics, *Materials Today: Proceedings* 80 2023; 2642–2646.
- [10] K. Shaheed, A. Mao, I. Qureshi, M. Kumar, Q. Abbas, I. Ullah and X. Zhang, A systematic review on physiological-based biometric recognition systems: Current and future trends, *Archives of Computational Methods in Engineering*. 2021; 1–44.
- [11] N. Bala, R. Gupta and A. Kumar, Multimodal biometric system based on fusion techniques: A review, *Information Security Journal: A Global Perspective*. 2022; 31(3): 289–337.
- [12] A.N. Uwaechia and D.A. Ramli, A comprehensive survey on ECG signals as new biometric modality for human authentication: Recent advances and future challenges, *IEEE Access* 9. 2021; 97760–97802.
- [13] Chen, Z. and Zhang, J., Gao, J., Li, P. A survey on deep learning for multimodal data fusion. *Neural Computation*, 2020; 32(5): 829-864.
- [14] Kurdthongmee, Wattanapong, and Piyadhida Kurdthongmee. "Fast and accurate pupil estimation through semantic segmentation fine-tuning on a shallow convolutional backbone, *HighTech Innov.*" (2024): 447-461.
- [15] Vladimir Zh. Kuklin¹, Naur Z. Ivanov¹, Alexander N. Muranov¹, Islam A. Alexandrov¹, Elena Yu. Linskaya, "Improving the Reliability of Biometric Authentication Processes Using a Model for Reducing Data Drift", *Emerging Science Journal*, 2024, Vol. 8, No. 6.
- [16] Adebayo, Bakare Mustaphaa, Kalaierasi Sonai Muthu Anbananthen, Saravanan Muthaiyah, and Saravanan Nathan Lurudusamy. "Comparative Analysis of Deep Learning Models for Part of Speech Tagging in the Malay Language." *HighTech and Innovation Journal* 5, no. 2 (2024): 272-281.
- [17] Bowyer KW, Chang K, Flynn P. A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Computer vision and image understanding*. 2006 Jan 1;101(1):1-5.
- [18] Poh N, Bengio S. Why do multi-stream, multi-band and multi-modal approaches work on biometric user authentication tasks. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing 2004 May 17; 5: 893*.
- [19] Ross A, Jain AK. Multimodal biometrics: An overview. In *2004 12th European signal processing IEEE conference 2004 Sep 6;1221-1224*.
- [20] Zhang X, Cheng D, Jia P, Dai Y, Xu X. An efficient android-based multimodal biometric authentication system with face and voice. *IEEE Access*. 2020 Jun 1; 8: 102757-72.
- [21] Li, S. Z. Jain AK and Kumar A. Biometric recognition: an overview. *Second generation biometrics: The ethical, legal and social context*. 2012 Mar; 6: 49-79.
- [22] Chen, X., Flynn, P. J., and Bowyer, K. W. (2005, October). Fully automated facial symmetry axis detection in frontal color images. In *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)* (pp. 106-111). IEEE.
- [23] Nguyen PT, Huynh VD, Vo KD, Phan PT, Elhoseny M, Le DN. Deep Learning based Optimal Multimodal Fusion Framework for Intrusion Detection Systems for Healthcare Data. *Computers, Materials & Continua*. 2021 Mar 1;66(3).
- [24] Eskandari M, Toygar Ö, Demirel H. Feature extractor selection for face-iris multimodal recognition. *Signal, image and video processing*. 2014 Sep;8(6):1189-98.
- [25] Mahalingam G, Ricanek K. LBP-based periocular recognition on challenging face datasets. *EURASIP Journal on Image and Video processing*. 2013 Dec;2013:1-3.
- [26] Lu Y, Khan M, Ansari MD. Face recognition algorithm based on stack denoising and self-encoding LBP. *Journal of intelligent systems*. 2022 Apr 21;31(1):501-10.
- [27] Ghayoumi M. A review of multimodal biometric systems: Fusion methods and their applications.

In 2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS) 2015 Jun 28 (pp. 131-136). IEEE.

[28] Nigam I, Vatsa M, Singh R. Ocular biometrics: A survey of modalities and fusion approaches. *Information Fusion*. 2015 Nov 1;26:1-35.

[29] Ramachandram D, Lisicki M, Shields TJ, Amer MR, Taylor GW. Bayesian optimization on graph-structured search spaces: Optimizing deep multimodal fusion architectures. *Neurocomputing*. 2018 Jul 12; 298:80-9.

[30] Mei Y, Lan T, Imani M, Subramaniam S. A Bayesian optimization framework for finding local optima in expensive multimodal functions. In *ECAI 2023* 2023 (pp. 1704-1711). IOS Press.

[31] Bianco S, Napoletano P. Biometric recognition using multimodal physiological signals. *IEEE access*. 2019 Jun 19; 7: 83581-8.

[32] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kegl, Algorithms for hyper-parameter optimization, in *Proc. Adv. Neural Inf. Process. Syst.*, 2011; 2546–2554.

[33] J. Bergstra, D. Yamins, and D. D. Cox, “Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures,” in *Proc. 30th Int. Conf. Int. Conf. Mach. Learn.*, 2013, pp. 1-115–1-123

[34] F. Hutter, H. H. Hoos, and K. Leyton-Brown, Sequential model-based optimization for general algorithm configuration,” in *Proc. Int. Conf. Learn. Intell. Optim. Springer*, 2011;507–523.

[35] Byeon H, Raina V, Sandhu M, Shabaz M, Keshta I, Soni M, Matrouk K, Singh PP, Lakshmi TV. Artificial intelligence-enabled deep learning model for

multimodal biometric fusion. *Multimedia Tools and Applications*. 2024 Oct;83(33):80105-28.

[36] Coelho KK, Tristão ET, Nogueira M, Vieira AB, Nacif JA. Multimodal biometric authentication method by federated learning. *Biomedical Signal Processing and Control*. 2023 Aug 1;85:105022.

[37] El_Rahman S, Alluhaidan AS. Enhanced multimodal biometric recognition systems based on deep learning and traditional methods in smart environments. *Plos one*. 2024 Feb 15;19(2):e0291084.

[38] Salturk S, Kahraman N. Deep learning-powered multimodal biometric authentication: integrating dynamic signatures and facial data for enhanced online security. *Neural Computing and Applications*. 2024 Jul;36(19):11311-22.

[39] Shivthare SK, Sharma YK, Patil RD. Multimodal Biometric in Computer Vision. *Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision*. 2023 Dec 1:1-29.

[40] Vieira J, Leitingner E, Sarajlic M, Li X, Tufvesson F. Deep convolutional neural networks for massive MIMO fingerprint-based positioning. In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) 2017 Oct 8:1-6.

[41] KumaaraMohanta, Mohapatra S. Development of Multimodal Biometric Framework for Smartphone Authentication System. *International Journal of Computer Applications*. 2014 Sep;102(7):6–11. DOI: 10.5120/17825-8597.

[42] Soleymani M, Dabouei A, Taherkhani F, Iranmanesh SM, Dawson J, Nasrabadi NM. Quality-Aware Multimodal Biometric Recognition. *IEEE Transactions on Biometrics, Behavior, and Identity Science*. 2022 Jan;4(1):97–116.