Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 22 March 2026
Revised 28 April 2026
Accepted 27 May 2026

Natural Resources for Human Health 2026, Vol. 6 (2s): 710–721

KEYWORDS: Biometric, Chimp Optimization Algorithm, feature level fusion, Gray-Level Co-Occurrence Matrix, Grey Wolf Optimizer, Smart healthcare

A Metaheuristic-Optimized Deep Neural Network for Feature-Level Fusion in Multimodal Biometrics for Secure Healthcare Systems

Bharathi R[1], Dr. M B Anandaraju[2]

[1]Research Scholar, ECE, BGS Institute of Technology, Adichunchanagiri University, Mandya, Karnataka, India, 571448.
[2]Professor, ECE, BGS Institute of Technology, Adichunchanagiri University, Mandya, Karnataka, India, 571448.
bharathir@bgsit.ac.in

ABSTRACT: Biometric security in healthcare systems is quickly becoming an important issue in the data security industry. An enormous obstacle remains the development of multimodal biometrics systems for smart environments that improve detection rates and accuracy. Traditional approaches using varying degrees of traits fusion are contrasted with MBS that use Convolutional Neural Networks (CNNs). This paper details the creation of a multimodal biometric identification system that enhances the accuracy and security of the system by fusing fingerprint and iris scanning at the feature level in smart healthcare systems. For feature extraction, it makes use of the Gray-Level Co-Occurrence Matrix (GLCM), Histogram of Oriented Gradients (HOG), and Local Binary Pattern (LBP). The Chimp Optimization Algorithm (ChOA) feature optimizer and Grey Wolf Optimizer (GWO) are utilized for feature selection. Using a Resnet, Densenet, and suggested Deep Neural Network (DNN) is the last stage in the categorization process. The results indicate that the DNN with ChOA outperformed DNN with GWO optimizer, with an accuracy rate of 98.14 %, AUC OF 0.992, F1-Score of 91.2 % and kappa with 0.959. This evidence is believed to using DNN with ChOA optimizer is an optimum choice for multimodal biometric system in smart healthcare applications.

1. INTRODUCTION

To increase the precision and security of recognition in smart healthcare system, this research suggested hybrid multimodal biometric systems (MBS) that combine various biometric characters. Using quaternion algorithms and RKHS-based dimensionality reduction, they presented three methods for deep feature fusion of five biometric characteristics [1]. The third approach combined features using deep learning. Compared to uni-biometric and other multi-modal systems, experimental findings from six databases shown that deep fusion significantly improves accuracy, strengthens the system against spoof attacks, and boosts security. Since authentication deals with proving the identity of the sender, receiver, logical server software, or device, it may be seen as a security function [2]. In light of the rational evolution of international standards, new technologies are being implemented in stages, and various identification techniques have been proposed for the transmission of safety-relevant information. Among these technical uses, biometrics stands out as the most practical since it can instantly and correctly identify a person based on a physiological characteristic. Biometric capabilities that place an emphasis on native two-factor authentication include voice recognition, iris, fingerprint, and face recognition. However, there is a potential privacy concern with tracking physiological templates for preservation and capture in the event that they leak. The attempts to secure data sovereignty are hampered by the fact that different countries have different legal stances. The inability to cancel transactions in the event of a compromise, slow enrollment rates, and environmental factors affecting sensor accuracy are all examples of technological difficulties that need attention [3]. Centralized matching services provide a security risk since they are susceptible to hacking. One of the latest methods incorporates AI into identification via characteristics including liveness detection, template quality evaluation, and improved face recognition. The second is that future identity verification will be more robust since deep learning models take user behaviors into account. The key issue with biometrics-intelligent automation interfaces, however, is the lack of clarity over how to do this while protecting people's privacy. One of the other areas that needs fixing is the fact that various health care system solutions don't have the same specs and can't work together [4].

2. LITERATURE SURVEY

A study on an MBS that combines features of the finger vein and surface was presented. The system is based on Near-Infra-Red (NIR) images of the fingers [5]. Finally, they are able to identify someone by fusing their biometric data using a fuzzy method. Researchers ran experiments on data from three different databases: NIRHI, HKPU, and UTFVP (University of Twente Finger Vein Pattern). The Lin-ear Binary Pattern is used to extract features for Finger Texture, which are subsequently classified using Support Vector Machines [6]. Finger Vein makes use of transfer learning using CNNs that have already been trained, namely VGG16, AlexNet, and VGG19. One method involves optimizing the intensity of the images before preprocessing them, and the other involves preprocessing near-infrared (NIR) images prior to CNN training. When compared to HKPU and UTFVP, NIRHI performs better in both the unimodal and multimodal configurations. With an astounding identification accuracy, the suggested multimodal system outperforms the most current state-of-the-art systems, which achieved high accuracy [7].

Provide the best approach for feature fusion in MBS using the Gravitational Search Algorithm (GSA) [8]. The idea is to use a combination of different biometric features to make recognition more accurate. They compare their method's performance to that of conventional GSA and test it at several threshold settings [9]. Their solution outperforms previous methods on real-world datasets in terms of recognition accuracy and computing efficiency, and it reaches the maximum accuracy at specific thresholds, too [10].

In this work, the authors provide a multimodal fusion architecture that makes use of DNN for feature-level, score-level, and pixel-level fusion [11]; they suggest smart fusion techniques such as class-1 and modality assessment; and they optimize the design using spatial, intensity, and channel fusion approaches [12]. In order to preprocess and extract features from iris, finger knuckle, palm, finger vein, and fingerprint images, the ERMOTMBA method with SU-NBO is presented in paper. Prior to using EC models like as DBN, CNN, and Bi-LSTM optimized by the SU-NBO algorithm, it employs median filtering and a variety of feature extraction techniques [13]. There has been some investigation into the possibility of using Cancelable CNN (CCNN) to secure multimodal biometric data, specifically focusing on iris and fingerprint biometric features. For extra safety, features are combined into one vector, CCNN is used to reduce them, and finally, a user-supplied seed is multiplied [14].

3. METHODOLOGY

Developing a recognition system in smart healthcare systems that can successfully identify user identities is the aim of this research. As seen in figure 1, this work uses Contrast Limited Adaptive Histogram Equalization (CLAHE) pre-processing, feature extraction using LBP, GLCM, and HOG, feature optimization, and classification using DNN to achieve this aim. DNNs are a type of ANN network's which are tailored to analyze pixel data and are commonly utilized for image recognition and processing.

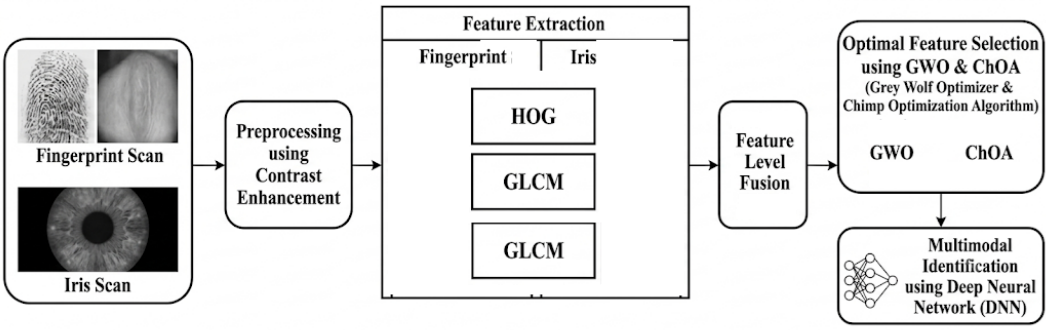

Figure 1: DNN model for multimodal biometric system

3.1 Multimodal Data

Fifty people provided biometric photos, including fingerprint and iris modalities. The multimodal biometric dataset of Person 1 utilised in this study, as shown in Figure 2 Following data augmentation, 1000 photos were created and utilised as a separate data set for this study. To produce a complete and balanced dataset, samples were obtained from both the left and right sides of each individual.

Figure 2: Biometric Traits of Person1 from own data set for multimodal

3.2 Pre-processing

Improving recognition accuracy in MBS requires pre-processing iris and fingerprint pictures. At first, every characteristic is scaled to a consistent 125x125 size. To emphasise minute details, fingerprint pictures are separated, ridge orientations are calculated, noise is minimised, and ridges are thinned and binarized. By employing CLAHE before to feature extraction, these pre-processing procedures enhance feature extraction and system resilience under various circumstances.

3.3 Feature Extraction

The GLCM, HOG and LBP are the feature extraction techniques used in biometric recognition, including iris, and fingerprint recognition. By obtaining significant and unique information from biometric samples (finger and iris), these methods are essential to multimodal biometric identification systems.

3.3.1 Gray-Level Co-Occurrence Matrix (GLCM)

For a given gray level image I with gray levels ranging from 0 to L-1, the GLCM P(i,j) is calculated as follows:

$$P(i,j) = \sum_{(x,y) \in \Omega} \delta(I(x,y) = i, I(x+\delta_x, y+\delta_y) = j) \quad (1)$$

In this equation, 'δ' represents the relationship offset (distance and direction), (x,y) are the pixel coordinates, and 'Ω' represents the image's total pixel coordinates. Using four different orientations (0°, 45°, 90°, and 135°), the GLCM approach efficiently captures the textures and patterns in the photos by extracting 88 texture characteristics. Each orientation contributes 22 features.

3.3.2 Local Binary Patterns (LBP)

LBP is a texture descriptor that labels pixels based on their neighborhood and is robust against changes in lighting. One is assigned to a neighboring pixel if its gray value is greater than the central pixel; otherwise, a zero is assigned. Equations (2) and (3) determine the LBP code value for each pixel. In the proposed work, 236 features are extracted using LBP for an image.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)\, 2^p \quad (2)$$

$$s(x) = \begin{cases} 1, x \geq 0; \\ 0, otherwise \end{cases} \quad (3)$$

3.3.3 Histogram of Oriented Gradients (HOG)

HOG is a sophisticated feature extraction approach commonly used in biometric security systems for iris, and fingerprint recognition. The approach accentuates the image's structure by recording the distribution of gradient orientations. HOG breaks each image into tiny cells and calculates a gradient orientation histogram for each one. In the present research 324 features per each image have been extracted with gradient and edge structure orientation values.

3.4 Feature Fusion of Fingerprint, Ear and Iris

Feature-level fusion by union was employed, combining feature vectors of iris (F_iris) as in equation (4) and ght Fingerprint feature (F_fingerprint) vectors as in equation (5) into a single fused vector as in equation (6). This higher-dimensional vectors are used to enhance system reliability and to improve system identification accuracy.

$$\text{F\_iris} = F_I = [F_{I1}, F_{I2}, F_{I3}, \ldots\ldots. F_{In}] \tag{4}$$

$$\text{F\_fingerprint} = F_{fp} = [F_{fp1}, F_{fp2}, F_{fp3} \ldots\ldots\ldots F_{fpn} \tag{5}$$

$$\text{Fused feature} = \text{F\_iris} \;||\; \text{F\_fingerprint} \tag{6}$$

Here, a combination of two traits is fused to obtain a final vector with length of 648 features combining 88 features from GLCM, 324 from HOG and 236 for LBP for each modality. The fused vectors are then saved in the database during the training phase.

3.5 Feature Optimization Algorithm for Multimodal Biometric System

The complete recognition pipeline is expressed as two co-operating procedures. Algorithm 1 covers the first stage, in which every enrolled iris and fingerprint sample is enhanced, described by the three texture and gradient descriptors, and fused at feature level into a single normalised representation. Algorithm 2 covers the second stage, in which the fused representation is compressed by a binary metaheuristic wrapper, using either the Grey Wolf Optimizer or the Chimp Optimization Algorithm, and the retained subset is used to train and evaluate the proposed deep neural network.

Algorithm 1: Stage-1 - Preprocessing, Multi-Descriptor Feature Extraction and Feature-Level Fusion

Input: Iris image set $\{I_{i,k}^{iris}\}$ and fingerprint image set $\{I_{i,k}^{fp}\}$ for $i = 1, \ldots, N$ subjects and $k = 1, \ldots, S$ samples per subject; CLAHE clip limit c and tile grid $t \times t$; target spatial size 125×125.
Output: Normalised fused feature matrix $Z_{norm} \in \mathbb{R}^{M \times D}$ with $M = N \cdot S$ and $D = 1296$; identity label vector $y \in \{1, \ldots, N\}^M$.

$m \leftarrow 0$
for $i = 1$ **to** N **do**
  for $k = 1$ **to** S **do**
    for each modality $\mu \in \{iris,\ fp\}$ **do**
      Resize and convert to grey scale: $\tilde{I}^\mu = \mathcal{G}\left(\mathcal{R}_{125\times125}(I_{i,k}^\mu)\right)$
      Contrast enhancement: $\hat{I}^\mu = CLAHE_{c,t}(\tilde{I}^\mu)$
      GLCM texture descriptor over $\theta \in \{0°, 45°, 90°, 135°\}$:
        $P_\theta(u, v) = \sum_{(x,y)\in\Omega} \delta\left(\hat{I}^\mu(x, y) = u,\ \hat{I}^\mu(x + \delta_x, y + \delta_y) = v\right)$
        $x_\mu^{GLCM} = [f_1, \ldots, f_{22}]_{\theta=0°} \| \cdots \| [f_1, \ldots, f_{22}]_{\theta=135°} \in \mathbb{R}^{88}$
      HOG orientation descriptor: $x_\mu^{HOG} \in \mathbb{R}^{324}$
      LBP texture code for every pixel, with $s(x) = 1$ if $x \geq 0$ and $s(x) = 0$ otherwise:
        $LBP_{P,R} = \sum_{p=0}^{P-1} s\left(g_p - g_c\right) 2^p, \qquad x_\mu^{LBP} \in \mathbb{R}^{236}$
      Modality descriptor by serial concatenation:
        $x_\mu = \left[x_\mu^{GLCM} \,\|\, x_\mu^{HOG} \,\|\, x_\mu^{LBP}\right] \in \mathbb{R}^{648}$
    end for
    Feature-level fusion by union of the two modality descriptors:
      $z = \left[x_{iris} \,\|\, x_{fp}\right] \in \mathbb{R}^D, \qquad D = 648 + 648 = 1296$
    $m \leftarrow m + 1; \quad Z(m, :) \leftarrow z^\top; \quad y(m) \leftarrow i$
  end for
end for
for $j = 1$ **to** D **do**
  Column statistics: $\mu_j = \frac{1}{M} \sum_{m=1}^{M} Z(m, j)$ and $\sigma_j = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \left(Z(m, j) - \mu_j\right)^2}$
  Z-score normalisation: $Z_{norm}(m, j) = \frac{Z(m,j) - \mu_j}{\sigma_j + \epsilon}, \quad m = 1, \ldots, M$
end for
return $Z_{norm},\ y$

Natural Resources for Human Health
eISSN: 2583-1194

Algorithm 2: Stage-2 - Binary Metaheuristic Feature Optimisation (GWO / ChOA) and DNN Classification

Input: Normalised fused matrix $Z_{norm} \in \mathbb{R}^{M \times D}$ and labels y from Algorithm 1; population size NP; maximum iteration count T_{max}; optimiser mode $\in \{GWO,\ ChOA\}$; fitness trade-off weight $\alpha \in (0,1)$; number of cross-validation folds F.
Output: Best binary selection mask b^*, optimised feature matrix Z_{opt}, trained network $\mathcal{N}^*$ and performance metrics.

1: **for** $p = 1$ **to** NP **do**

2: Initialise the continuous position $x_p \sim \mathcal{U}(0,1)^D$

3: Map the position to a binary mask through the sigmoid transfer function:

$$T(x_{p,j}) = \frac{1}{1+e^{-x_{p,j}}}, \qquad b_{p,j} = \begin{cases} 1, & T(x_{p,j}) \geq r,\ r \sim \mathcal{U}(0,1) \\ 0, & \text{otherwise} \end{cases}$$

4: Evaluate the wrapper fitness, in which E_{DNN} is the F-fold cross-validated classification error and $\|b_p\|_1$ is the number of selected features:

$$f(b_p) = \alpha \cdot E_{DNN}(Z_{norm}[:, b_p = 1], y) + (1-\alpha) \cdot \frac{\|b_p\|_1}{D}$$

5: **end for**

6: Rank the population by $f(\cdot)$ and initialise the social hierarchy:
GWO leaders $x_\alpha, x_\beta, x_\delta$ (three fittest); ChOA leaders $x_{Att}, x_{Bar}, x_{Cha}, x_{Dri}$ (four fittest)

7: **for** $t = 1$ **to** T_{max} **do**

8: Linearly decaying exploration coefficient: $a = 2 - 2 \cdot \frac{t}{T_{max}}$

9: **for** $p = 1$ **to** NP **do**

10: Draw $r_1, r_2 \sim \mathcal{U}(0,1)$ and form $A = 2a \cdot r_1 - a$ and $C = 2r_2$

11: **if** mode $= GWO$ **then**

$$D_\ell = |C_\ell \cdot x_\ell - x_p|, \quad X_\ell = x_\ell - A_\ell \cdot D_\ell, \quad \ell \in \{\alpha, \beta, \delta\}$$
$$x_p \leftarrow \frac{X_\alpha + X_\beta + X_\delta}{3}$$

12: **else if** mode $= ChOA$ **then**

Generate the chaotic driving coefficient m_t from the chaotic map

$$d_\ell = |C_\ell \cdot x_\ell - m_t \cdot x_p|, \quad X_\ell = x_\ell - A_\ell \cdot d_\ell, \quad \ell \in \{Att, Bar, Cha, Dri\}$$
$$x_p \leftarrow \frac{X_{Att} + X_{Bar} + X_{Cha} + X_{Dri}}{4}$$

13: **end if**

14: Re-map the updated position to a binary mask: $b_{p,j} = \begin{cases} 1, & \frac{1}{1+e^{-x_{p,j}}} \geq r \\ 0, & \text{otherwise} \end{cases}$

15: **if** $\|b_p\|_1 = 0$ **then** set one randomly chosen bit of b_p to unity **end if**

16: Re-evaluate $f(b_p)$ with the fitness function of step 4

17: **end for**

18: Update the hierarchy leaders from the current fitness ranking

19: **end for**

20: Retain the globally best mask $b^* = \mathrm{argmin}_b f(b)$

21: Build the optimised subset $Z_{opt} = Z_{norm}[:, b^* = 1] \in \mathbb{R}^{M \times D^*}$, where $D^* = \|b^*\|_1$

22: **for** $l = 1$ **to** L **do** (forward propagation of the DNN, with $a^{(0)} = z_{opt}$)

23: $z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)}, \qquad a^{(l)} = \varphi(z^{(l)})$

24: **end for**

25: Soft-max identity posterior: $\hat{y}_n = \frac{e^{z_n^{(L)}}}{\sum_{q=1}^{N} e^{z_q^{(L)}}}$

26: Minimise the categorical cross-entropy loss by back-propagation:

$$\mathcal{L} = -\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} y_{m,n} \log \hat{y}_{m,n}, \qquad W^{(l)} \leftarrow W^{(l)} - \eta \frac{\partial \mathcal{L}}{\partial W^{(l)}}$$

27: $\mathcal{N}^* \leftarrow$ trained network; compute accuracy, precision, recall, F1-score, kappa, AUC, FPR and EER on the held-out partition

28: **return** $b^*, Z_{opt}, \mathcal{N}^*$ and the metric set

3.6 DNN

Deep Neural Networks (DNNs) are a kind of ANN that, because to its multi-layer architecture connecting input and output layers, are able to simulate complex data correlations and patterns. Data delivered from input to output layers in a simple DNN may be processed by many hidden layers. In order to confirm a person's identity, different neural network (DNN) multimodal biometric identification systems employ a variety of biometric modalities, such as fingerprint and iris pictures. Figure 3 shows the design of a DNN for MBS. A DNN in a multimodal biometric identification system would employ a variety of inputs, including images of the user's fingerprints, knuckles, and palm, to provide an estimate of the user's identity. To train the DNN, one may use a large dataset of biometric information, where each data point could be a distinct individual and where data could come from a variety of biometric modalities. For a multimodal biometric identification system to be effective, it is necessary to extract discriminative and suggestive characteristics from each biometric modality. The last step is to feed the DNN the best recovered features and instruct it to make the association between the characteristics and the person's identify. Using the DNN to forecast biometric data that has never been observed before is possible after training. Additionally, it may be used with matching algorithms and other algorithms to form a robust and accurate system for multimodal biometric identification. One may mathematically represent each layer in a DNN using Eq. (7) and (8): where 1 is the current layer index and z^ is the layer's position. The variables ((l)) and "W" represent the weighted sum of the inputs at layer 1, "a^((1-1))" is the activation output from the previous layer, "b^((1))" is the bias vector for layer 1, and "F(. is the activation function," for example ReLU, sigmoid, or softmax.

$$Z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)} \tag{7}$$

$$a^{(l)} = \sigma\left(z^{(l)}\right) \tag{8}$$

For classification tasks, the output layer might use a softmax activation to produce probabilities:

$$\hat{y}_i = \frac{\exp(z_i)}{\sum_{j=1}^{k} \exp\left(z_j\right)} \tag{9}$$

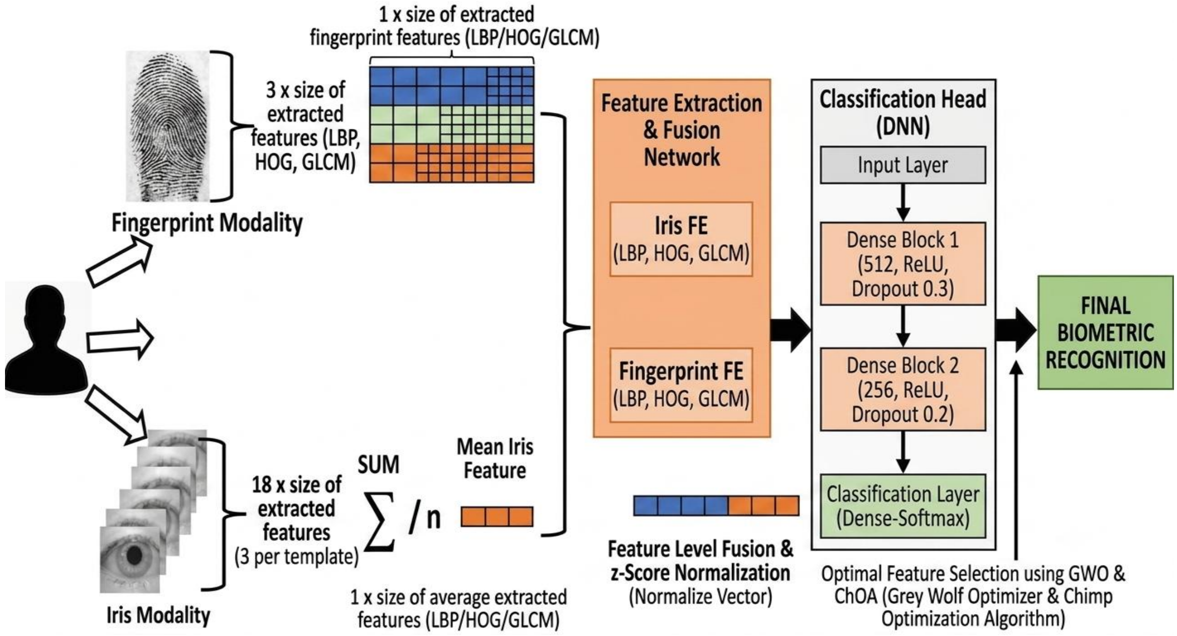

Figure 3: DNN model for multimodal biometric system

3.7 Performance evaluation

An Optimum Multimodal Biometric System for Iris, Fingerprint, and Ear Recognition employs feature-level fusion to merge retrieved information into a single vector, therefore increasing accuracy and robustness [15]. Performance is measured using accuracy, precision, recall, and F1-score, determined using the following equations [16]:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FN+FP}} \tag{10}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{11}$$

$$\text{F1 Score} = \frac{2*\text{Recall}*\text{Precision}}{\text{Precision}+\text{Recall}} \quad (12)$$

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

4 RESULTS AND DISCUSSION

The performance of the proposed approach is evaluated with the datasets of fingerprints and iris from own data set collected in the university with the approval form the ethical committee and from the individuals without misuse of data other than this research. In the proposed approach, handcrafted descriptors are used to extract the features across all biometric traits. The features extracted from each trait are concatenated, scaled, and passed through a DNN classification head to perform the final recognition. Here, Resent 50, Densenet -121 and proposed DNN models have been trained. The author conducted the experiment in MATLAB R2026B, with Operating System: Windows 10, Processor: Intel i7, Memory: 16 GB, and Graphics: GeForce GTX 950 GPU.

Two pre-trained models - Resent 50, Dense net 121 and a proposed DNN were fine-tuned with and without optimization for MBS. The DL models are trained using fingerprint and iris photos from 1000 photographs, with 80% of the images utilized for training and 20% for testing. Tables 1 and 2 show the results of Resent 50, Densenet -121 models that obtain the best accuracy with the least amount of loss during training and validation. In the scenario of concatenation fusion, the top accuracy achieved by ResNet50 was 84.14, 86.50 and 88.42 without optimization and using GWO, ChOA optimizers respectively. Similarly, Dense Net 121 was achieved an accuracy of 87.25, 89.57 and 91.21 respectively. When compared to ResNet50, Dense Net 121 achieves better accuracy. When it comes to feature-level fusion, Dense net with ChOA to be the optimal multimodal fusion method.

Table 1: Resent for multimodal features

Fold	Accuracy	Precision	Recall	F1-Score	AUC	Specificity	FPR	MCC	Kappa
Without Optimization	84.14	83.87	84.33	84.05	0.885	83.91	0.161	0.682	0.681
GWO	86.50	86.21	86.80	86.52	0.901	86.35	0.137	0.730	0.729
ChOA	88.42	88.15	88.62	88.35	0.912	88.24	0.118	0.768	0.767

Table 2: Dense net for multimodal features

Fold	Accuracy	Precision	Recall	F1-Score	AUC	Specificity	FPR	MCC	Kappa
Without Optimization	87.25	86.91	87.59	87.24	0.908	87.42	0.135	0.744	0.743
GWO	89.57	89.23	89.81	89.50	0.925	89.39	0.107	0.790	0.789
ChOA	91.21	90.96	91.52	91.24	0.938	91.08	0.091	0.824	0.823

Deep Neural Network (DNN)

A DNN was further refined in this research. For the purpose of multimodal application, a five-layer DNN was used as a DL approach. Table 3 provides the architecture of the suggested DNN that was employed in the investigation with validation, using GWO and ChOA optimizers to assess the model performance.

eISSN: 2583-1194

Table 3: Proposed Model with Parameters

DNN	Parameters
Image Dimension	128 × 128
Dense Layer	8 neurons
Activation Function	ReLU
Dense Layer	16 neurons
Activation Function	ReLU
Dense Layer	32 neurons
Activation Function	ReLU
Dense Layer	48 neurons
Activation Function	ReLU
Dense Layer	64 neurons
Output Layer	Softmax Function

One measure of a model's performance is its Mean Average Precision (mAP). The definition of Mean Average Precision varies between systems. A standard mAP curve produced by training and validating the DNN model is shown in Figure 4.

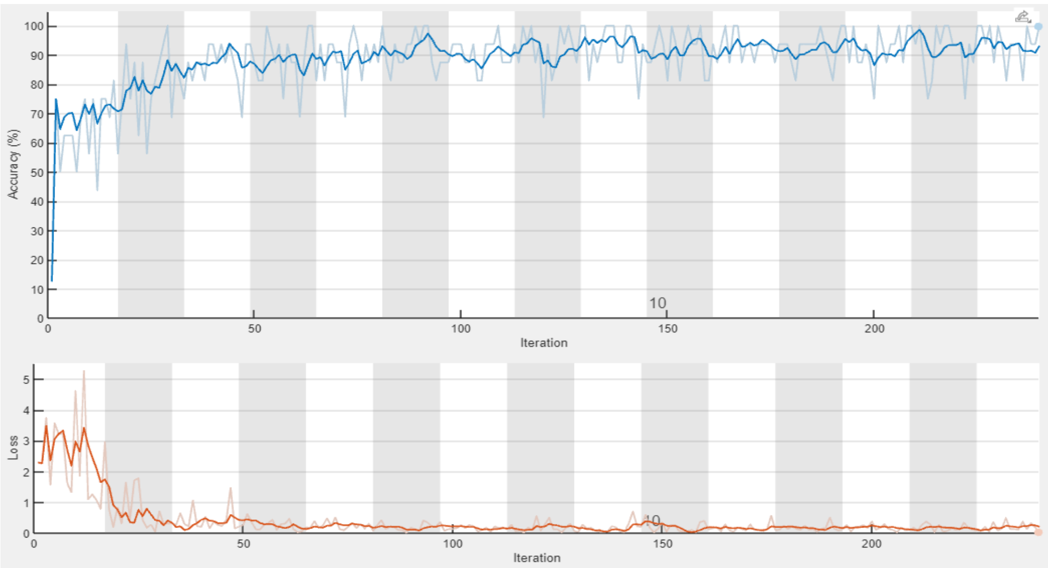

Figure 4: mAP curve for the proposed DNN with ChOA model for multimodal biometric system

Figure 5 shows a parallel coordinates diagram of the hyperparameters. Optimal parameters may be seen in the graphic, which shows all optimization attempts performed by the Bayesian optimizer across different numbers, datasets, and matchers. Keep in mind that most orientation and cell hyperparameter values are inside the sweet spot.

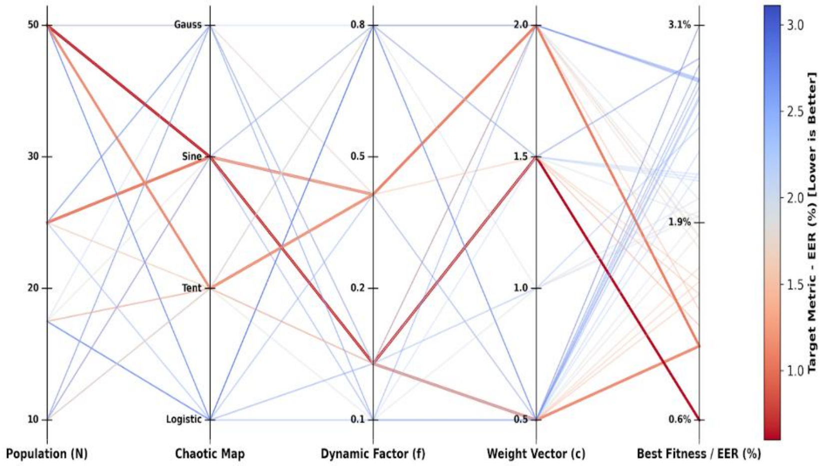

Figure 5: Optimized DNN with ChOA model for multimodal biometric system

The trained DNN was fine-tuned for use in a multimodal biometric system, both with and without optimization. The results of the DNN model that achieves the highest accuracy with the lowest loss during training and validation are shown in Tables 4. Maximum DNN accuracy in the concatenation fusion scenario was 93.83% without optimization, 96.51% with GWO, and 98.14% with ChOA optimizers, respectively. Figure 6 highlights that, accuracy achieved by proposed DNN is superior than that of ResNet50 and Dense Net 121. A low equal error rate (EER) has been observed in the DNN with ChOA optimizer. The best multimodal fusion approach is the one that uses Proposed DNN with ChOA for feature-level fusion.

Table 4: DNN for multimodal features

Fold	Accuracy	Precision	Recall	F1-Score	AUC	Specificity	FPR	MCC	Kappa	EER
Without Optimization	93.83	93.50	94.21	93.75	0.952	93.65	0.064	0.876	0.875	4.8
GWO	96.51	96.23	96.8	96.51	0.98	96.40	0.036	0.93	0.929	2.0
ChOA	98.14	97.81	98.23	98.26	0.992	97.91	0.021	0.96	0.959	0.8

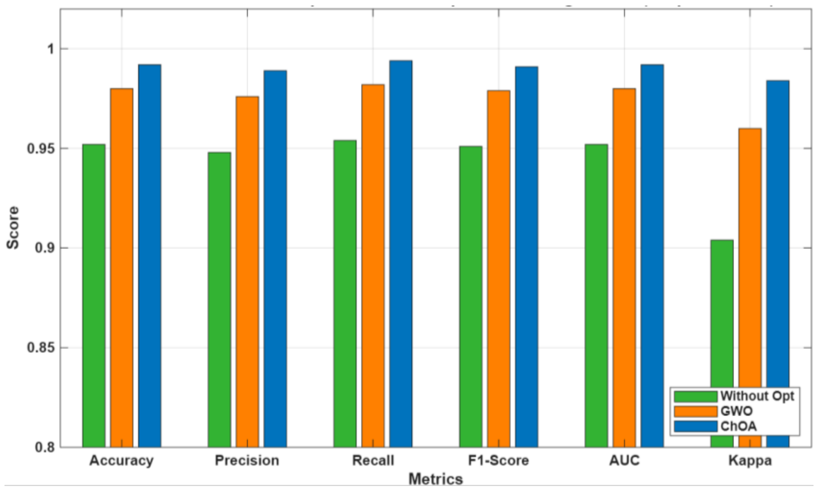

Figure 6: Performance metrics comparison across optimization algorithms

Natural Resources for Human Health
eISSN: 2583-1194

Figure 7 shows the Equal Error Rate (EER) at 0. 8 for the proposed DNN with ChOA. A more nuanced balance between the rates of erroneous acceptance and false rejection seems to be implied here. One benefit of feature-level fusion is a more accurate and robust multimodal biometric system, and the enhanced EER is a prime example of this.

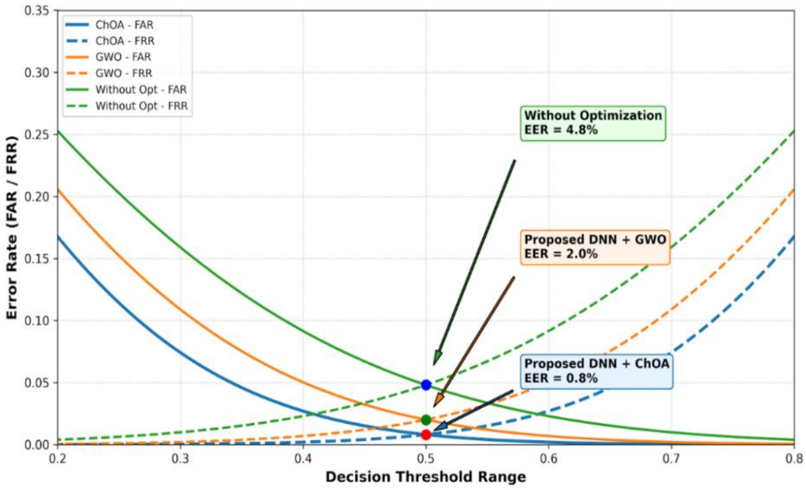

Figure 7: EER for proposed DNN model

The ROC analysis shown in figure 8 is a clear progression in performance across the three models. ResNet-50 begins with an AUC of 0.885 without optimization, improves to 0.901 with GWO, and reaches 0.912 with ChOA. DenseNet-121 performs better overall, starting at 0.908, rising to 0.925 `with GWO, and topping out at 0.938 with ChOA. The proposed DNN leads throughout, moving from a strong baseline of 0.952 to 0.980 with GWO, and finally achieving the highest discriminative score of 0.992 with ChOA.

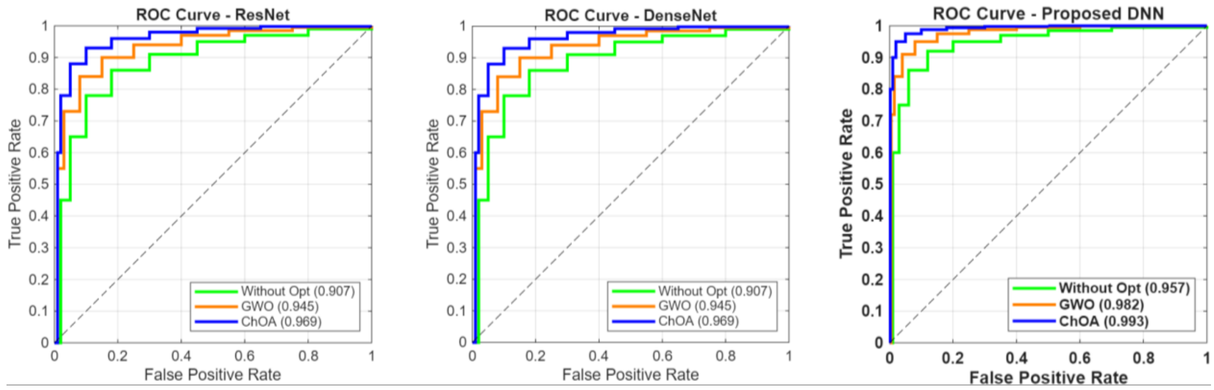

Figure 8: Comparison of ROC plot for multimodal biometric system with and with optimizer

Table 5 presents a comparative analysis of proposed model with state-of-the-art models, demonstrating a moderate improvement in accuracy over existing approaches.

Table 5: Comparison of proposed Model with state of art of models

Sl. No	Authors	Methods used	Accuracy in %
1	Ammour et al. [17]	Adaptive CNN & ResNet	95.67
2	Khatri et al. [18]	CNN-based model	96.15
3	Edwards et al. [19]	Extreme Learning Machine + PSO	97.14
4	Zhai et al. [20]	Alexnet	85.60
5	Proposed Model optimized DNN		98.14

5. CONCLUSION

A multimodal biometric system for in smart healthcare systems has been developed using deep learning models. Reliable features are produced for database matching when DNN models are used to extract unique features. Two optimizers were applied to the data using Resnet 50, dense net 121 and proposed DNN. Multimodal identification biometric systems that employ two traits, including the palm print trait and iris. The results of the investigation indicate that for the multimodal biometric applications, the proposed DNN performs better than the Resnet 50 and dense net 121 using with and without optimization approach. DNN approach achieves the greatest accuracy of all the methods in the suggested system with a highest accuracy of 93.83 % without optimizer and 98.14% with ChOA optimizer. The low FPR of 0.021 and average kappa of 0.950 show good ability while reducing false alarms. Additionally, the model's excellent generalization ability is shown by its balanced metric data, recall of 98.23, AUC 99.20, F1-score of 98.26 and EER 0.8, making it perfect for reliable multimodal biometric system. This high accuracy shows how well the system reduces false acceptance and rejection rates in multimodal recognition systems. It also highlights the potential of DNN model to advance biometric authentication technologies and suggests future optimization and performance improvements under various conditions.

REFERENCES

[1] D. Jagadiswary and D. Saraswady, “Biometric authentication using fused multimodal biometric,” Procedia Computer Science, vol. 85, pp. 109–116, 2016, doi: 10.1016/j.procs.2016.05.187.

[2] N. Ajay, H. S. Mohan, I. S. Rajesh, and S. Gupta, “A deep learning and blockchain-based security framework for cloud data protection,” Journal of Discrete Mathematical Sciences and Cryptography, vol. 29, no. 6, pp. 2553–2587, 2026, doi: 10.47974/JDMSC-2823.

[3] I. Boucherit, M. O. Zmirli, H. Hentabli, and B. A. Rosdi, “Finger vein identification using deeply-fused convolutional neural network,” Journal of King Saud University – Computer and Information Sciences, vol. 34, no. 3, pp. 646–656, Mar. 2022, doi: 10.1016/j.jksuci.2020.04.002.

[4] K. Gunasekaran, J. Raja, and R. Pitchai, “Deep multimodal biometric recognition using contourlet derivative weighted rank fusion with human face, fingerprint and iris images,” Automatika: Journal for Control, Measurement, Electronics, Computing and Communications, vol. 60, no. 3, pp. 253–265, 2019, doi: 10.1080/00051144.2019.1565681.

[5] Z. Tao, X. Zhou, Z. Xu, S. Lin, Y. Hu, and T. Wei, “Finger-vein recognition using bidirectional feature extraction and transfer learning,” Mathematical Problems in Engineering, vol. 2021, Art. no. 6664809, pp. 1–11, 2021, doi: 10.1155/2021/6664809.

[6] H. M. Therar, E. A. Mohammed, and A. J. Ali, “Multibiometric system for iris recognition based convolutional neural network and transfer learning,” IOP Conference Series: Materials Science and Engineering, vol. 1105, no. 1, Art. no. 012032, 2021, doi: 10.1088/1757-899X/1105/1/012032.

[7] R. M. Jasmine, J. Jasper, and M. R. Geetha, “An efficient secure cryptosystem using improved identity based encryption with multimodal biometric authentication and authorization in cloud environments,” Wireless Networks, vol. 31, no. 1, pp. 545–565, Jan. 2025, doi: 10.1007/s11276-024-03780-8.

[8] Y. Ma, Z. Huang, X. Wang, and K. Huang, “An overview of multimodal biometrics using the face and ear,” Mathematical Problems in Engineering, vol. 2020, Art. no. 6802905, pp. 1–17, 2020, doi: 10.1155/2020/6802905.

[9] S. A. Abdulrahman and B. Alhayani, “A comprehensive survey on the biometric systems based on physiological and behavioural characteristics,” Materials Today: Proceedings, vol. 80, pp. 2642–2646, 2023, doi: 10.1016/j.matpr.2021.07.005.

[10] S. Gupta, I. S. Rajesh, C. Singh, U. N. Ranjitha, K. Siddesha, and P. K. Sekharamantry, “A hybrid CEFL approach using gradient sparsification and quantization for medical imaging,” International Journal of Computational Intelligence in Engineering, vol. 1, no. 1, pp. 1–15, 2026.

[11] N. Bala, R. Gupta, and A. Kumar, “Multimodal biometric system based on fusion techniques: a review,” Information Security Journal: A Global Perspective, vol. 31, no. 3, pp. 289–337, 2022, doi: 10.1080/19393555.2021.1974130.

[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, vol. 9, pp. 97760–97802, 2021, doi: 10.1109/ACCESS.2021.3095248.

[13] J. Gao, P. Li, Z. Chen, and J. Zhang, “A survey on deep learning for multimodal data fusion,” Neural Computation, vol. 32, no. 5, pp. 829–864, May 2020, doi: 10.1162/neco_a_01273.

[14] 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, vol. 14, no. 8, Art. no. 222, 2022, doi: 10.3390/fi14080222.

[15] B. Abd El-Rahiem, F. E. Abd El-Samie, and M. Amin, "Multimodal biometric authentication based on deep fusion of electrocardiogram (ECG) and finger vein," Multimedia Systems, vol. 28, no. 4, pp. 1325–1337, 2022, doi: 10.1007/s00530-021-00810-9.

[16] S. Dodge, J. Mounsef, and L. Karam, "Unconstrained ear recognition using deep neural networks," IET Biometrics, vol. 7, no. 3, pp. 207–214, May 2018, doi: 10.1049/iet-bmt.2017.0208.

[17] N. Ammour, Y. Bazi, and N. Alajlan, "Multimodal approach for enhancing biometric authentication," Journal of Imaging, vol. 9, no. 9, Art. no. 168, 2023, doi: 10.3390/jimaging9090168.

[18] M. Khatri and A. Sharma, "Deep learning approach based on iris, face, and palmprint fusion for multimodal biometric recognition system," International Journal of Performability Engineering, vol. 19, no. 6, pp. 407–416, 2023, doi: 10.23940/ijpe.23.06.p6.407416.

[19] T. Edwards and M. S. Hossain, "Effectiveness of deep learning on serial fusion based biometric systems," IEEE Transactions on Artificial Intelligence, vol. 2, no. 1, pp. 28–41, Feb. 2021, doi: 10.1109/TAI.2021.3064003.

[20] Y. Zhai et al., "A novel finger-knuckle-print recognition based on batch-normalized CNN," in Proceedings of the 13th Chinese Conference on Biometric Recognition (CCBR 2018), Lecture Notes in Computer Science, vol. 10996. Cham, Switzerland: Springer, 2018, pp. 11–21, doi: 10.1007/978-3-319-97909-0_2.