Biometric Identification Advances: Unimodal to Multimodal Fusion of Face, Palm, and Iris Features

Ola Najah KADHIM[1,2], Mohammed Hasan ABDULAMEER[3]
[i]*Department of Computer Science, Faculty of Computer Science & Mathematics, University of Kufa, Najaf, 54001, Iraq*
[2]*Technical Institute of Al-Mussaib, Al-Furat Al-Awsat Technical University, Najaf, 54001, Iraq*
[3]*Department of Computer Science, Faculty of Education for Girls, University of Kufa, Najaf, 54001 Iraq*
ola.najah@atu.edu.iq

***Abstract*—Due to increased information security concerns, biometric recognition technology has become more important. Unimodal biometrics still work effectively, but they struggle with noise sensitivity and spoof attack susceptibility since they rely on a single data source. This paper uses advances in deep learning and machine learning to propose new unimodal systems for the palm, face, and iris. These models use deep wavelet transform networks (WTN) for face and iris identification and deep convolutional neural networks (CNNs) for palmprint identification. In addition, we introduce a novel multimodal biometric system based on unimodal systems. We get 98.29% for face, 98.86% for palmprint, and 95.59% for iris in individual unimodal systems with Support Vector Machines (SVM). This is done by using the new property MULB dataset, which has many biometric features. The multimodal system achieves 99.88% accuracy and a 0.0186 equal error rate, underscoring the relevance of several biometric features and the superior performance of the identification system.**

***Index Terms*—deep learning, feature extraction, feature level fusion, multimodal biometrics identification, machine learning.**

I. INTRODUCTION

The demand for precise user recognition systems that can control access to these breakthroughs has increased, consequently, the rapid advancement of contemporary technology. Biometric recognition technologies stand out among the alternatives as the most sophisticated and efficient solution. A scientific subject called biometrics uses physical characteristics like the iris and face as well as distinctive behavioral qualities like voice and gait to identify people using automated methods [1-2]. Because biometric data is distinct and resistant to loss, theft, or copying, there are numerous advantages associated with this approach in comparison to conventional authentication methods such as passwords [3].

Biometric systems can be divided into single biometric and multimodal biometric categories based on the number of identifying modalities employed. Even though unimodal systems are reliable, effective, and better than traditional methods, they do have some problems, such as noise, universality, being easy to spoof, and similarities between people [4]. Multimodal biometric systems integrate many modalities to improve recognition accuracy and reliability in response to these difficulties. These methods overcome the limitations of individual features and offer a more reliable authentication procedure by merging a variety of traits [5].

There exist various categories of multimodal biometric systems, encompassing multi-sensor, multi-sample, multi-algorithm, multi-instance, and multimodal approaches [6]. In order to integrate information from several modalities, these systems use various fusion-level techniques. These techniques include sensor-level fusion, feature-level fusion, match-score-level fusion, and decision-level fusion [7]. In the field of biometric feature extraction and recognition [8], researchers widely apply machine learning (ML) techniques, including classifiers and feature extraction from raw data.

However, there are limitations on how well ML approaches can differentiate and choose characteristics across many domains. Researchers created deep learning (DL), a more recent branch of machine learning, to overcome these restrictions. In order to extract low-level information and translate it into abstract features, DL uses artificial neural networks (ANNs) with several hidden layers [9]. This research carefully examines the performance of unimodal systems for face, palm, and iris attributes independently. Then, it merges them into a multimodal system utilizing feature-level fusions in light of the excellent performance of deep learning approaches in multiple recognition tasks. For palmprint recognition, deep convolutional neural networks (CNNs) are used, while face and iris recognition utilize deep wavelet transform networks (WTNs). Several factors influence the decision to use face, palm, and iris patterns when building a multimodal biometric system. Its dependability and security are increased by the iris's constant and unwavering texture throughout a person's life. The iris is one of the most precise biometrics currently accessible, noted for its exact recognition despite its intrusiveness [10]. Because it is simple to use and seems natural, facial recognition is a good tool for recognizing people in images. The consistency and distinctive qualities of the palmprint modality also offer useful information for identifying people [11]. Face, palmprint, and iris patterns work best together as a comprehensive and reliable multimodal biometric system because of these characteristics. In this study, we introduce unique unimodal recognition methods for iris, face, and palm recognition. These systems benefit from the latest advancements in deep learning and machine learning methodologies. The systems we suggest make use of deep wavelet transform networks (WTN) for recognizing faces

Digital Object Identifier 10.4316/AECE.2024.01010

and irises, as well as deep convolutional neural networks (CNNs) for recognizing palmprints. We seamlessly incorporate feature extraction and classification procedures into these models, utilizing a wide range of machine-learning classifiers to increase their efficacy.

Then, in order to attain an exceptional level of performance, we combine those unimodal systems with a novel multimodal biometric system. There are several classification techniques utilized, including SoftMax, support vector machines (SVM) with one versus. one (OvO), logistic regression (LR) with one vs. rest (OvR), K-nearest neighbor's (KNN), Random Forest (RF), and Naive Bayes (NB). This study also looks at fusion strategies in the context of a multimodal system. It focuses on feature-level fusion using a range of methods, including concatenation, the sum rule, the weighted sum rule, the mean, the standard deviation, and the median. The next lines give a summary of the paper's major accomplishments:

- The introduction of a unique homogeneous multimodal biometric dataset of features found in the face, palm, and iris;
- Utilizing advanced deep learning approaches to construct distinctive unimodal biometric systems for iris, palm, and facial identification;
- The development of a novel multimodal biometric technique that fuses feature-level data from the iris, palm, and face features;
- Conducting several trials on the unique dataset and comprehensively evaluating the suggested models.

II. Related works

Many research publications have developed unimodal and multimodal systems based on various biometric recognition technologies, including face, palm, iris, and others. This section gives an overview of current research that has used multimodal and unimodal machine learning systems based on conventional machine learning and deep learning methods.

Bachay and Abdulameer [12], building on their earlier research, took on the difficulty of palmprint authentication, highlighting the need for effective feature extraction and authentication methodologies. They bring out a hybrid autoencoder (AE) and convolutional neural network (CNN) model to solve these problems. Although the hybrid technique has promise, the study does not critically evaluate any potential drawbacks or determine if the model is successful in overcoming the particular obstacles involved in palmprint recognition. Muntadhar and Abdulameer [13] conducted a study on face recognition and created an intelligent face recognition system with a focus on addressing problems related to aging. They used deep learning in their strategy, specifically convolutional neural networks (CNNs). With recognition accuracy rates of 98.7% on the Age dataset and 99.4 % on the FG-Net dataset, their method impressively demonstrated the power of deep learning in addressing face recognition difficulties, particularly those related to ageing.

The same authors examined a number of difficulties with facial recognition, including stance, lighting, occlusion, and ageing, in similar research [14]. They used convolutional neural network (CNN) and support vector classifier (SVC) models. Their approach included pre-processing, feature extraction using a pre-trained CNN variation called the VGG face model, and classification using the support vector machine (SVM). Astonishingly, their method produced a maximum accuracy rate of 97.00 % for all age groups in the dataset, demonstrating the effectiveness of CNN and SVC in addressing age-related differences in facial recognition tasks. However, unimodal systems are commonly favored over multimodal biometric systems. However, multimodal biometric systems are commonly favored over unimodal systems for improved security. This is because they have the capacity to incorporate various biometric qualities, decrease vulnerability to spoof attacks, and raise total authentication accuracy. Many researchers have investigated different multimodal biometric techniques; nevertheless, a thorough analysis of their methodologies indicates certain shortcomings and possibilities for development.

In the beginning of the studies that used multimodal biometrics, Hariton *et al.* [15] presented three biometric techniques for person verification: fingerprints, face recognition, and voice recognition. It assesses their accuracy using a majority voting technique. The study shows 92 % recognition precision, suggesting potential for future complex biometric cards. Still, more testing across larger databases is required. For instance, Micucci and Iula [16] enhanced performance by fusing 3D palmprint and hand geometry characteristics. However, the assessment mainly concentrates on the PolyU database, which might not accurately reflect several real-world problems, including position, lighting, and occlusions.

A remarkable multimodal hand biometric system incorporating Finger Knuckle Print (FKP) and palmprint data was provided by Attia *et al.* [17]. However, the lack of a larger dataset in the reported results may constrain the generalizability of their method. The hybrid model by Arjun and Prakash [18] that combines decision and feature-level fusion exhibits improved recognition rates. Still, a more thorough examination of its performance under various circumstances would paint a better image of its resilience. The multimodal face, finger vein, and iris recognition system developed by Alay and Al-Baity [19] achieves good accuracy. However, the assessment only uses data from the SDUMLA-HMT dataset, which only partially represents the difficulties experienced in real-world situations. The method by Mahmoud *et al.* [20] that combines iris and face traits shows promise in terms of recognition accuracy and fusion time reduction. Still, more testing across other databases is required to confirm its efficacy in more general scenarios. Last but not least, Ammour *et al*. [10] face-iris multimodal system shows outstanding recognition rates but primarily focuses on two distinct databases, necessitating study in more vast and varied datasets to determine its robustness. While these studies demonstrate the potential of multimodal biometric approaches, it is crucial to take into account the dataset diversity constraints and the requirement for homogenous biometric types in order to ensure their practical effectiveness in a variety of security applications, which is what we are attempting to demonstrate in this paper.

III. The Proposed Uni-Model and Multi Model System

The present section gives a comprehensive review of unimodal biometric identification systems as well as multimodal biometric systems.

A. Face Identification System

In this section, we analyze the facial identification system depicted in Fig. 1, which heavily relies on WTN. Our suggested technique consists of three main steps in this face-based unimodal system: (1) pre-processing the facial image; (2) feature extraction using the WTN model; and (3) classification using the SoftMax classifier. During the pre-processing stage, we first use multi-task cascaded convolutional neural networks (MTCNN) to identify faces in the image [21]. MTCNN is designed for face detection and is based on deep convolutional neural networks. Each of the three consecutive neural networks (P-Net, R-Net, and O-Net) is dedicated to a different aspect of face detection.

These networks work together to precisely locate and recognize faces in the image. To ensure consistent facial identification across scales, we enlarge the original image to several sizes and create an image pyramid before using these networks. We next resize all of the images to the proper dimensions, crop the face image sections, and finally convert all of the input images into grayscale representations. Fig. 2 displays the samples obtained after completing the pre-processing stage.

Following this, using the deep wavelet transform network (WTN) [22], the procedure involves the integration of a series of wavelet transformations, followed by the application of a modulus operator, in order to extract the distinctive characteristics from the pre-processed facial image. A deep wavelet has numerous layers, each of which uses the input from the previous layer as its output. According to Fig. 3, each layer consists of three procedures.

An image denoted as '*I*' will be considered. The initial scattering coefficient represents the image's average; the calculation can be achieved through the process of convoluting the image with a scaling filter. The initial scattering coefficient represents the image's average, which can be calculated by convolving the image with a scaling filter (low-pass filter) 'φ_j', as shown in (1) [23]:

$$S_{0,j}(I) = I * \varphi_j \tag{1}$$

where '*j*' represents a specific scale. The initial-layer scattering coefficients are acquired through convolution applied to the input image '*I*' using a wavelet filter '$\psi_{\lambda 1}$' at a particular scale '*j*', followed by taking the modulus of the resultant coefficients. Subsequently, these coefficients undergo low-pass filtering using a scaling filter 'φ_j'. The scattering coefficients of the initial-layer $S_{1,j}(\lambda_1, I)$ can be obtained by (2) [23]:

$$S_{1,j}(\lambda_1, I) = \left|I * \psi_{\lambda 1}\right| * \varphi_j(I) \tag{2}$$

Through a multi-step process, we derive the secondary-layer scattering coefficients. First, convolve the image 'I' with a wavelet filter '$\psi_{\lambda 1}$' at a specific scale '*j*', and calculate the modulus of the resulting coefficients.

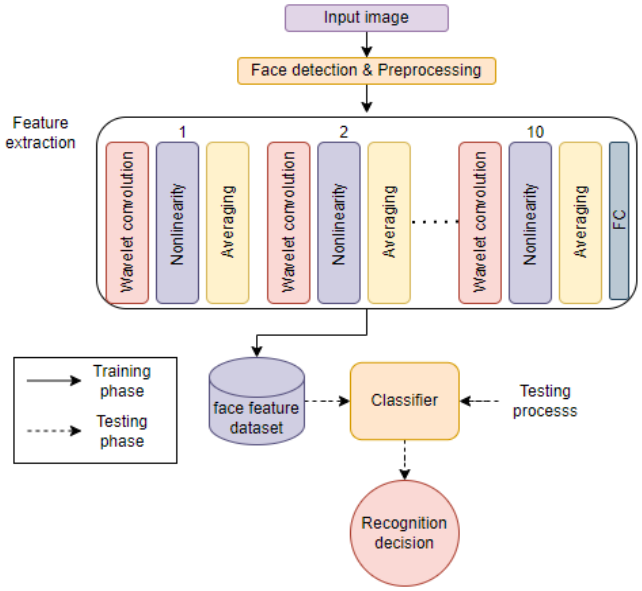

Figure 1. Face identification system

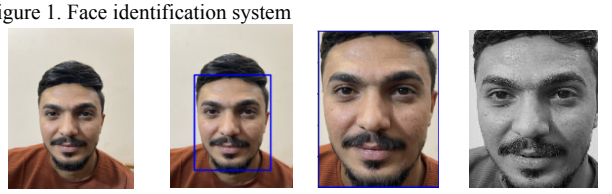

Figure 2. Face pre-processing steps

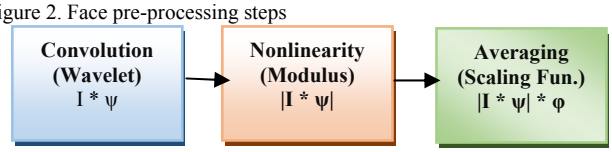

Figure 3. Operations of Wavelet scatter

Then, these modulus values undergo another convolution, this time with a different wavelet filter '$\psi_{\lambda 2}$' at a distinct scale, followed by taking the modulus again. A scaling filter 'φ_j' is then used to apply a low-pass filtering operation. the scattering coefficients of the second-layer $S_{2,j}((\lambda_1, \lambda_2), I)$ can be obtained by (3) [23]:

$$S_{2,j}\left((\lambda_1, \lambda_2), \mathrm{I}\right) = \left\|\mathrm{I} * \psi_{\lambda_1}\right| * \psi_{\lambda 2}\left| * \varphi_j(\mathrm{I}) \right. \tag{3}$$

Note that the extent of local translation invariance is defined by the parameter '*j*', which represents the width of the low-pass filter. Beyond the number of scales that the transformation can generate, 'φ' symbolizes a low-pass filter, 'ψ' represents a wavelet, and 'λ' pertains to the rotation operations. Since '$S_{1,j}$', and '$S_{2,j}$' corresponds to the outputs of low-pass filters, they can be down-sampled based on the filter's width raised to the power of '2^j'. The last step in the wavelet scattering network is the average pooling layer, which figures out the average value of each scattering coefficient over the space domain. The final output of the network consists of a vector that aggregates all calculated averages, and, as a result, the proposed method successfully captures both the low-level and high-level textural elements included in the input signal or image. In (4) [23] determines the output for this third layer:

$$S_{k,j}((\lambda_1, \lambda_2, ..., \lambda_k), I) = \left\|I * \psi_{\lambda 1}\right| * ... * \psi_{\lambda k}\left| * \varphi_j(I) \right. \tag{4}$$

In this context, '*S*' denotes the ultimate output vector, whereas '*k*' signifies the total number of layers within the

network. The values λ_1, λ_2,…,λ_k represent the potential angles for the orientation of the wavelet filters at each layer. Additionally, '*j*' values assigned to each layer dictate the range of scales that the transformation can generate. The output vector '*S*' that is produced provides vital information on the local spatial arrangement of texture features. It is an invaluable tool for a variety of image analysis and classification tasks. The Deep Wavelet Transform Network used in this process consists of three layers, each with ten nodes. A set of filters that closely mimic the convolution filters employed in convolutional neural networks for image processing at each layer.

Next, the fully connected layer applies the Rectifier Linear Unit (ReLU) activation function. By creating connections between each neuron in the previous layer and each neuron in the following layer, this layer helps the network understand non-linear combinations of information. A group of feature vectors is produced at this stage, which proves helpful in the classification procedure. After this stage, the SoftMax and SVM classifiers predict the labels for the input patterns.

B. Palmprint Identification System

This section goes into depth regarding the proposed palmprint identification system, as can be observed in Fig. 4. Our suggested palmprint system approach consists of three independent phases: (1) initial palmprint image pre-processing; (2) feature extraction using a CNN model; and (3) classification using SoftMax. We take a progressive strategy throughout the pre-processing step. We begin by identifying the region of interest (ROI) within the palmprint images. Following that, we resize all images to their proper proportions before converting all input images to grayscale representations. The approach in [12] recovers the vital core part of the palmprint image. The approach in [12] determines the ROI as a square shape and then converts it into a 192×192-pixel image.

Fig. 5 depicts all of the pre-processing processes. The steps stated in [12] are as follows:

- Using a Gaussian smoothing procedure on the input image;
- Binarizing the smoothed image using a threshold, represented as '*T*';
- Obtaining the binary image's boundaries using a boundary-tracking technique;
- Determining points positioned between fingers to determine the 2D ROI pattern inside the border image;
- Finally, we extract the region of interest (ROI).

Following this, a powerful convolutional neural network (CNN) model extracts characteristic from the palmprint's Region of Interest (ROI). This CNN model used in our study has nine layers in total, including three convolutional layers, three max-pooling layers, one flattens layer, one fully connected layer (FC), and a dropout layer.

This architectural cascade starts with a convolutional operation in the first layer. In each convolution layer that follows, a Rectifier Linear Unit (ReLU) activation function is used. During these early phases, we keep the image proportions at 200×200 pixels. After that, a Max-Pooling layer is used to downsize the feature image to 100×100 pixels in size.

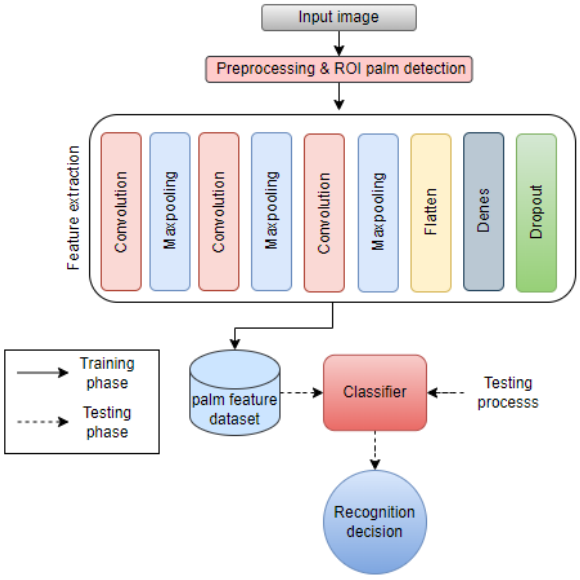

Figure 4. Palmprint identification system

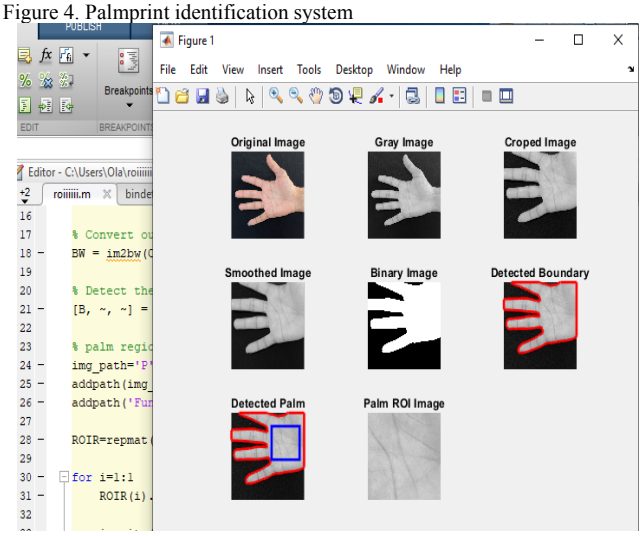

Figure 5. Palmprint pre-processing steps

The third layer, which serves as another convolutional layer, maintains an output size similar to the second convolutional layer. After that, we add a Max-Pooling layer, which results in an output with dimensions of 50×50 pixels. The following layer continues as a convolution layer, resulting in a 50×50-pixel image. The Max-Pooling layer further reduces the output to 25×25 pixels. A critical flatten layer is introduced, which is in charge of converting the feature map into a vector representation. The fully connected (FC) layer is next added, with its unit count dynamically adjusted depending on the preceding layer and the number of categories under consideration. To minimize network complexity and avoid overfitting risks, we carefully introduce a dropout layer with a dropout rate of 0.5 before the classifier. We optimize the CNN model using the Adam optimization approach and train it using the categorical cross-entropy loss function. After this sophisticated feature extraction procedure, the SoftMax classifier predicts the labels associated with the input patterns.

C. Iris Identification System

This section delves into the iris identification algorithm, which employs WTN as a feature extractor. The approach described before for facial identification, as seen in Fig. 6, is quite similar to this technique. The strategy can be

delineated into three distinct stages: Pre-processing of iris images to maintain uniformity and standardizing the size of all iris images at this early step. Following that, we transform all of the incoming images to grayscale. We then subject the iris images to a series of meticulously staged techniques to extract distinct and distinctive features.

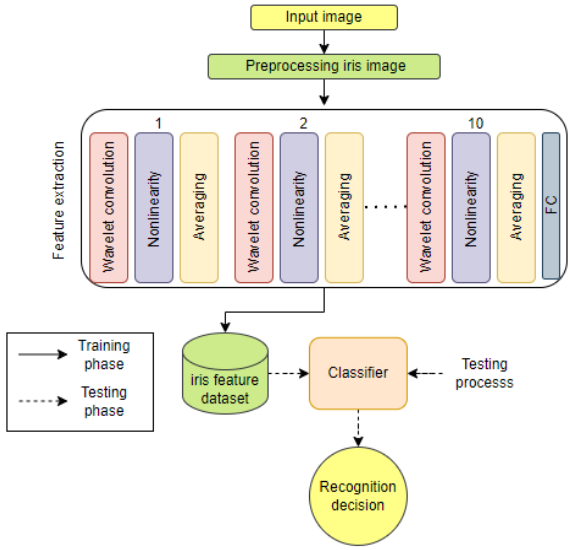

Figure 6. Iris identification system

The following operations comprise these steps: localize, segment, normalize, and extract features. Segmentation, in particular, refers to the task of splitting the image into individual pixels. Two circular borders are used to designate the inner and outer margins for segmenting the iris. An integro-differential operator and order statistics (maximum) are used to perform this segmentation, as shown in (5) [24].

$$\max\left(r,x_0,y_0\right)\left|G_\sigma\left(r\right)*\frac{\partial}{\partial_r}\int_{r,x_o,y_0}\frac{\mathrm{I}\left(x,y\right)}{2\pi r}ds\right| \quad (5)$$

In this system, 'I' represents the iris image, and '$I(x, y)$' means the pixel intensity at coordinates (x, y). '$G_\sigma(r)$' represents a Gaussian filter with radius 'r' and scale 'σ' applied radially. The sign 'r' represents the radius of several circular areas, each centred at (x_0, y_0). A '∂' denotes the standard deviation of the Gaussian distribution. The coordinates (x_0, y_0) are presumed to be the center of the iris. The parameters (r, x_0, y_0) specify the contour of a circle denoted by 'ds'. This integro-differential equation's purpose is to locate circle edges, a crucial step in image processing. In image processing, the operator scans the image domain (x_0, y_0) to identify the peak value within the blurred partial derivatives. This blurring is achieved through convolution with a smoothing function like a Gaussian with a scale of 'σ'. The operator searches for the highest derivative of the normalized contour integral of $I(x_0, y_0)$ along a circular arc 'ds' with radius 'r' centered at (x_0, y_0). This operator essentially acts as a circular edge detector that blurs with a scale determined by 'σ'. The process involves iteratively searching for the most significant contour integral derivative, progressively refining the analysis at finer scales. This search occurs across the parameter space of center coordinates and radius (x_0, y_0, r), defining a trajectory of contour integration. The normalization of the segmented iris involves representing it as a rectangular image. During normalization, we transform the segmented iris from Cartesian form (x, y) to polar form (r, θ), where 'r' ranges from 0 to 1, and 'θ' spans from 0 to 2π, representing the angle. Daugman's Rubber Sheet model [25]. During the normalization phase, the model effectively converts the circular texture of the iris into a corresponding rectangular shape. Fig. 7 illustrates the characteristic processing procedures of the iris.

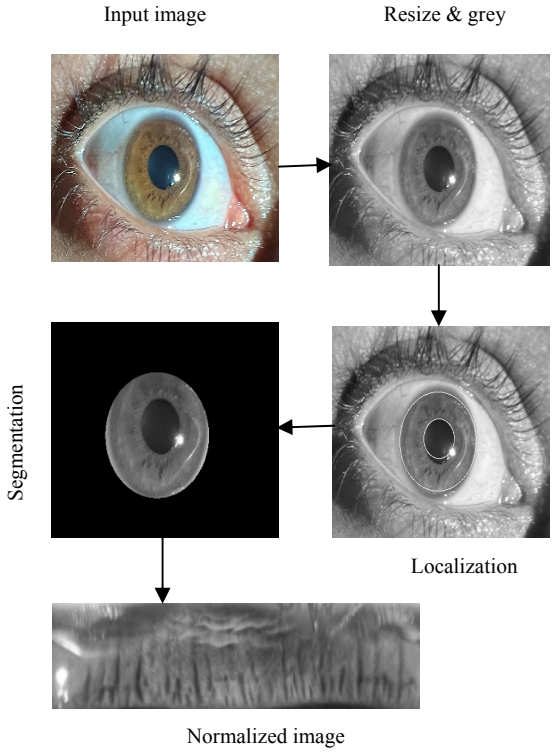

Figure 7. Iris pre-processing steps

Following this, the iris image that has been normalized is next subjected to the process of feature extraction via an advanced wavelet transform network (WTN) model similar to that described for face identification. The SoftMax classifier predicts the labels linked to the input patterns.

D. Multimodal Biometric System

Four core processes organize the system: face identification, palmprint identification, iris identification, and feature-level fusion. The multimodal system establishes the user's identification. It uses a deep wavelet transform network (WTN) to extract features from both facial and iris characteristics. Furthermore, it utilizes a deep convolutional neural network (CNN) to collect characteristics from palmprint traits. Finally, the model integrates the feature vectors from these three features and utilizes classification machine learning algorithms to detect and authenticate the user's identity. Fig. 8 depicts the general structure of the system using a block diagram.

IV. Features Level Fusion

Feature-level fusion refers to combining features obtained from various traits. The characteristics derived from the three separate traits are combined to produce new features that thoroughly reflect the user. The model learns to recognize and utilize these combined characteristics in the fusion strategy. Throughout the training phase, the model harmonizes the outputs of the fully connected layers of the

face, palmprint, and iris models. The aggregation process combines the feature vectors obtained from the fully connected layers of each of the three models into a unified vector. Researchers use various fusion procedures, such as the sum rule, weighted sum rule, concatenation, mean rule, standard deviation rule, and median rule.

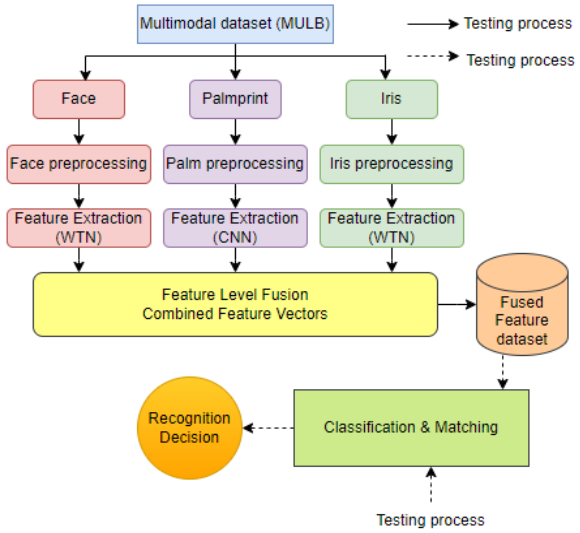

Figure 8. General block diagram of proposed multimodal

V. Classification Process

In order to strengthen the validity of our method, we used a variety of classification algorithms in our investigations. Support vector machine (SVM) with the One-vs-One (OvO) strategy, logistic regression (LR) with the One-vs-rest (OvR) strategy, K-nearest neighbors (KNN), random forest (RF), and Naive Bayes (NB) are among the approaches used. We subjected these approaches to a thorough comparison study to properly evaluate and confirm the efficacy and efficiency of our suggested methodology.

VI. Experimental Results

This section will present the data derived from the experiments done using the multimodal biometric system. We proposed a system that combines iris, palmprint, and facial traits for recognition. We will start by giving a summary of the novel datasets and the various parameters used in them for our experiments. Following that, we'll discuss the evaluation findings for each modality, including face, palm, and iris identification. We will now discuss the detailed assessment results of our proposed multimodal system, which merges face, palmprint, and iris recognition at the feature level. We use the accuracy rate, as in (6) [26], as our performance metric.

$$ACC = \frac{\left(T_{pos} + T_{neg}\right)}{\left(T_{pos} + T_{neg} + F_{pos} + F_{neg}\right)} \tag{6}$$

A crucial criterion for assessing and contrasting the effectiveness of biometric systems is the Equal Error Rate (EER). It is the point at which the false acceptance rate (FAR) and the false rejection rate (FRR) are equal. The false acceptance rate (FAR) is the percentage of impostor attempts (incorrectly accepting an unauthorized user). However, the false rejection rate (FRR) is the percentage of genuine attempts that are incorrectly rejected (failing to recognize an authorized user). The formula for FAR, FRR, and EER is shown in (7), (8), and (9) [27]:

$$FAR = \frac{F_{pos}}{\left(F_{pos} + T_{neg}\right)} \tag{7}$$

$$FRR = \frac{F_{neg}}{\left(T_{pos} + F_{neg}\right)} \tag{8}$$

$$EER = \frac{FAR + FRR}{2} \tag{9}$$

The probability of accurately identifying authorized users among all the examined scenarios is signified by the True Positive Rate (Tpos). The probability of correctly identifying authorized users who are not detected among all the tests run is indicated by the True Negative Rate (Tneg). The percentage of unauthorized users mistakenly discovered throughout all tests is known as the False Positive Rate (Fpos). In contrast, the False Negative Rate (Fneg) shows the proportion of unauthorized users that are mistakenly missed in all tests run. Our study separated every sub dataset in our new MULB dataset into training, validation, and testing subsets utilizing the subsequent ratios: (80:10:10), (60:20:20), and (70:10:20). The ratio (70:10:20) produced the best outcomes. As a result, in this study, we divided the data for each person into the following ratios: 70 % for training, 10 % for validation, and 20 % for testing.

A. MULB Dataset Description

To evaluate the suggested method, we used the novel multimodal biometric dataset MULB, which contains the corresponding biometric traits. A total of 176 people (118 men and 58 women) with ages ranging from 17 to 54 provided data for the three subsets of datasets that make up MULB: the face, palmprint, and iris datasets. 10,560 color images in the 'JPG' format make up the whole collection. Using the iPhone 14 Pro Max's micro camera, every biometric was painstakingly recorded. The Al-Furat Al-Awsat Technical University in Kufa, Iraq, produced the MULB dataset in the winter of 2023. The face dataset represents each individual with 20 images, showcasing a variety of stances, expressions, and accessories. There are 3,520 images of faces in all. The palmprint dataset consists of 20 images of the right hand of each individual taken at various angles. The iris dataset contains 20 images of the right iris for each individual captured under various lighting conditions and angles.

B. Evaluation of Unimodal Identification Systems

For each unimodal biometric, we used a SoftMax classifier to achieve accuracy rates of 97.87 %, 98.01 %, and 94.89 % for face, palmprint, and iris identification, respectively. We subsequently compared unimodal identification accuracy to popular machine learning classification techniques. We used the features from the model-based feature extractor we suggested in this evaluation. During the multiclass classification procedure, we grouped the test data into several class labels within the training data to generate model predictions. Tables I, II and III illustrate the comparison results for identification accuracy. In Table I, it is clear that the suggested model, when utilizing the SVM classifier for face identification, obtained the maximum accuracy, reaching 98.29 %. Table II

shows that the SVM classifier achieved a significantly higher accuracy of 98.86 % for palmprint identification compared to other classifiers. Table III shows that, when using the SVM classifier, the suggested model had the highest accuracy for iris identification (95.59 %).

TABLE I. A COMPARATIVE ANALYSIS OF FACE IDENTIFICATION ACCURACY ACROSS VARIOUS CLASSIFIERS

Classification Algorithms	Accuracy
K-nearest neighbors (KNN)	95.17
Naïve Bayes (NB)	95.45
Random forest (RF)	97.15
Logistic regression (LR) (OvR)	97.72
Support vector machine (SVM) (OvO)	**98.29**

TABLE II. A COMPARATIVE ANALYSIS OF ACCURACY IN PALMPRINT IDENTIFICATION ACROSS VARIOUS CLASSIFIERS

Classification Algorithms	Accuracy
K-nearest neighbors (KNN)	95.73
Naïve Bayes (NB)	95.02
Random forest (RF)	98.43
Logistic regression (LR) (OvR)	98.57
Support vector machine (SVM) (OvO)	**98.86**

TABLE III. A COMPARATIVE ANALYSIS OF IRIS IDENTIFICATION ACCURACY ACROSS VARIOUS CLASSIFIERS

Classification Algorithms	Accuracy
K-nearest neighbors (KNN)	87.78
Naïve Bayes (NB)	91.33
Random forest (RF)	93.89
Logistic regression (LR) (OvR)	95.31
Support vector machine (SVM) (OvO)	**95.59**

C. Evaluation of Multimodal Identification System

We performed multimodal biometric identification in the fourth experiment. Table IV shows the identification accuracies achieved by classifying the combined features using multiclass classification techniques.

TABLE IV. A COMPARATIVE ANALYSIS OF MULTIMODAL IDENTIFICATION ACCURACY ACROSS VARIOUS CLASSIFIERS AND FUSION TECHNIQUES

Classification algorithms	Feature fusion methods	Accuracy
K-nearest neighbors (KNN)	Sum rule	99.71
	Weighted sum rule	99.28
	concatenation	**99.85**
	Mean rule	99.71
	Std. rule	99.71
	Median rule	97.01
Naïve Bayes (NB)	Sum rule	93.18
	Weighted sum rule	93.18
	concatenation	**95.88**
	Mean rule	93.18
	Std. rule	93.18
	Median rule	82.10
Random forest (RF)	Sum rule	99.71
	Weighted sum rule	99.71
	concatenation	**99.85**
	Mean rule	**99.85**
	Std. rule	99.71
	Median rule	99
Logistic regression (LR) (OvR)	Sum rule	99.85
	Weighted sum rule	**99.88**
	concatenation	99.85
	Mean rule	99.85
	Std. rule	99.85
	Median rule	99.28
Support vector machine (SVM) (OvO)	Sum rule	99.85
	Weighted sum rule	**99.88**
	concatenation	99.85
	Mean rule	99.85
	Std. rule	99.85
	Median rule	99.14

It is noteworthy that when using the concatenation fusion approach, the K-nearest neighbors classifier attained the best accuracy of 99.85 %. Similar to this, using the concatenation fusion approach, the Naive Bayes classifier achieved the best accuracy of 95.88 %. Additionally, the concatenation fusion method enabled the Random Forest classifier to reach the same high accuracy of 99.85 %.

In contrast, when using the weighted sum fusion approach, both the SVM and the logistic regression classifiers were able to reach a perfect accuracy of 99.88 %. As a consequence, in both unimodal and multimodal identification scenarios, models based on logistic regression and SVM classifiers showed improved performance. Table V summarizes the Equal Error Rate (EER) values for the suggested biometric system. It is interesting that for multimodal identification situations, the EER rates for both the logistic regression and SVM classifiers were 0.0186. The EER of the proposed multimodal system was much lower than that of the separate unimodal biometric systems, which is significant.

TABLE V. EQUAL ERROR RATES FOR THE PROPOSED BIOMETRIC IDENTIFICATION METHODS

Biometrics	Classifier	EER
Face	K-nearest neighbors (KNN)	0.0674
	Naïve Bayes (NB)	0.0543
	Random forest (RF)	0.0472
	Logistic regression (LR) (OvR)	0.0419
	SVM (OvO)	**0.0327**
Palmprint	K-nearest neighbors (KNN)	0.0513
	Naïve Bayes (NB)	0.0668
	Random forest (RF)	0.0315
	Logistic regression (LR) (OvR)	0.0304
	SVM (OvO)	**0.0241**
Iris	K-nearest neighbors (KNN)	0.1985
	Naïve Bayes (NB)	0.0809
	Random forest (RF)	0.0734
	Logistic regression (LR) (OvR)	0.0578
	SVM (OvO)	**0.0519**
Multimodal	K-nearest neighbors (KNN)	0.0271
	Naïve Bayes (NB)	0.0791
	Random forest (RF)	0.0232
	Logistic regression (LR) (OvR)	**0.0186**
	SVM (OvO)	**0.0186**

VII. CONCLUSION

This work introduced three distinct unimodal biometric systems, each focusing on the identification of a single person using their face, palm, or iris, and each utilizing a different deep learning model. In addition, we proposed a cutting-edge multimodal biometric identification system that combines features from all three biometric qualities. The unimodal systems used a variety of machine learning classifiers in addition to WTN models for face and iris identification. The experimental procedure includes thorough preprocessing of the face, palmprint, and iris images used as input for the corresponding WTN and CNN models. We used the UMLB dataset, specifically created for feature extraction, to thoroughly train these models. We then used the collected features in a classification model for person identification, examining five distinct classifier types across the three modalities: face, palmprint, and iris identification. In the field of multimodal biometrics, SVM and logistic regression achieved a phenomenal 99.88% accuracy rate, making them the most successful classifiers. With face identification at 98.29 %, palmprint recognition at 98.86 %, and iris identification at 95.59 %, unimodal

biometrics achieved outstanding accuracy rates. With improved accuracy, fewer trainable parameters, and quicker training periods, our techniques outperformed previous systems. Notably, the multimodal system fared better in terms of equal error rate (EER) than individual unimodal biometric systems, proving its supremacy in person identification. In upcoming research projects, we will expand our multimodal system to include datasets that are subject to data consumption restrictions. We will investigate score-level fusion methodologies to further improve system performance.

ACKNOWLEDGMENT

Students and staff are thanked for their participation in creating and registering the multimodal biometric database used in this research.

REFERENCES

[1] Y. Wang, D. Shi, and W. Zhou, "Convolutional neural network approach based on multimodal biometric system with fusion of face and finger vein features," Sensors, vol. 22, no. 16, pp. 1-15, 2022. doi:10.3390/s22166039

[2] L. Huang and Y. Wu, "Structure-aware heatmap and boundary map regression based robust face alignment," Advances in Electrical Computer Engineering, vol. 23, no. 2, pp. 3-10, 2023. doi:10.4316/AECE.2023.02001

[3] R. Gad, A. El-Sayed, N. El-Fishawy, and M. Zorkany, "Multi-biometric systems: a state of the art survey and research directions," International Journal of Advanced Computer Science Applications, vol. 6, no. 6, pp. 128-138, 2015. doi:10.14569/IJACSA.2015.060618

[4] M. O. Oloyede and G. P. Hancke, "Unimodal and multimodal biometric sensing systems: A review," IEEE access, vol. 4, pp. 7532-7555, 2016. doi:10.1109/ACCESS.2016.2614720

[5] M. Singh, R. Singh, and A. Ross, "A comprehensive overview of biometric fusion," Information Fusion, vol. 52, no. 2, pp. 1-24, 2019. doi:10.1016/j.inffus.2018.12.003

[6] H. Al-Mahafzah, T. AbuKhalil, M. Alksasbeh, and B. Alqaralleh, "Multi-modal palm-print and hand-vein biometric recognition at sensor level fusion," International Journal of Electrical Computer Engineering, vol. 13, no. 2, pp. 1954-1963, 2023. doi:10.11591/ijece.v12i5

[7] M. I. Ahmad, W. L. Woo, and S. Dlay, "Non-stationary feature fusion of face and palmprint multimodal biometrics," Neurocomputing, vol. 177, pp. 49-61, 2016. doi:10.1016/j.neucom.2015.11.003

[8] V. Timcenko and S. Gajin, "Machine learning enhanced entropy-based network anomaly detection," Advances in Electrical Computer Engineering, vol. 21, no. 4, pp. 51-60, 2021. doi:10.4316/AECE.2021.04006

[9] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," Pattern Recognition Letters, vol. 141, pp. 61-67, 2021. doi:10.1016/j.patrec.2020.07.042

[10] B. Ammour, L. Boubchir, T. Bouden, and M. Ramdani, "Face–iris multimodal biometric identification system," Electronics, vol. 9, no. 1, pp. 1-18, 2020. doi:10.3390/electronics9010085

[11] M. Chaa, N.-E. Boukezzoula, and A. Attia, "Score-level fusion of two-dimensional and three-dimensional palmprint for personal recognition systems," Journal of Electronic Imaging, vol. 26, no. 1, pp. 1-12, 2017. doi:10.1117/1.JEI.26.1.013018

[12] F. M. Bachay and M. H. Abdulameer, "Hybrid deep learning model based on autoencoder and CNN for palmprint authentication," International Journal of Intelligent Engineering Systems, vol. 15, no. 3, pp. 488-499, 2022. doi:10.22266/ijies2022.0630.41

[13] M. H. Ibrahem and M. H. Abdulameer, "Age invariant face recognition model based on convolution neural network (CNN)," Journal of Al-Qadisiyah for Computer Science and Mathematics vol. 1, no. 1, pp. 96-110, 2023. doi:10.29304/jqcm.2023.15.1.1143

[14] M. H. Ibrahem and M. H. Abdulameer, "Age face invariant recognition model based on VGG face based DNN and support vector classifier," International Journal on Technical and Physical Problems of Engineering vol. 15, no. 45, pp. 232-240, 2023

[15] H. N. Costin, I. Ciocoiu, T. Barbu, and C. Rotariu, "Through biometric card in Romania: person identification by face, fingerprint and voice recognition," International Journal of Biological Medical Sciences, vol. 1, no. 4, pp. 264-269, 2006. doi:10.5281/zenodo.1071630

[16] M. Micucci and A. Iula, "Recognition performance analysis of a multimodal biometric system based on the fusion of 3D ultrasound hand-geometry and palmprint," Sensors, vol. 23, no. 7, pp. 1-14, 2023. doi:10.3390/s23073653

[17] A. Attia, S. Mazaa, Z. Akhtar, and Y. Chahir, "Deep learning-driven palmprint and finger knuckle pattern-based multimodal Person recognition system," Multimedia Tools Applications, vol. 81, no. 8, pp. 10961-10980, 2022. doi:10.1007/s11042-022-12384-3

[18] B. C. Arjun and H. N. Prakash, "Multimodal biometric recognition system using face and finger vein biometric traits with feature and decision level fusion techniques," International Journal of Computer Theory and Engineering, vol. 13, no. 4, pp. 123-128, 2021. doi:10.7763/IJCTE.2021.V13.1300

[19] N. Alay and H. H. Al-Baity, "Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits," Sensors, vol. 20, no. 19, pp. 1-17, 2020. doi:10.3390/s20195523

[20] R. O. Mahmoud, M. M. Selim, and O. A. Muhi, "Fusion time reduction of a feature level based multimodal biometric authentication system," International Journal of Sociotechnology Knowledge Development, vol. 12, no. 1, pp. 67-83, 2020. doi:10.4018/IJSKD.2020010104

[21] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE signal processing letters, vol. 23, no. 10, pp. 1499-1503, 2016. doi:10.1109/LSP.2016.2603342

[22] A. Rehman, M. Harouni, M. Omidiravesh, S. M. Fati, and S. A. Bahaj, "Finger vein authentication based on wavelet scattering networks," Computers, Materials Continua, vol. 72, no. 2, 2022. doi:10.32604/cmc.2022.016410

[23] J. Andén and S. Mallat, "Deep scattering spectrum," IEEE Transactions on Signal Processing, vol. 62, no. 16, pp. 4114-4128, 2014. doi:10.1109/TSP.2014.2326991

[24] R. T. Mohammed, H. Kaur, B. Alankar, and R. Chauhan, "Recognition of human Iris for biometric identification using Daugman's method," IET Biometrics, vol. 11, no. 4, pp. 304-313, 2022. doi:10.1049/bme2.12074

[25] A. Matin, F. Mahmud, S. T. Zuhori, and B. Sen, "Human iris as a biometric for identity verification," in Proc 2nd International Conference on Electrical, Computer & Telecommunication Engineering Bangladesh, 2016. doi:10.1109/ICECTE.2016.7879610

[26] S. Almabdy and L. Elrefaei, "Deep convolutional neural network-based approaches for face recognition," applied sciences, vol. 9, no. 20, pp. 1-21, 2019. doi:10.3390/app9204397

[27] H. M. L. Aung, C. Pluempitiwiriyawej, K. Hamamoto, and S. Wangsiripitak, "Multimodal biometrics recognition using a deep convolutional neural network with transfer learning in surveillance videos," computation, vol. 10, no. 127, pp. 1-15, 2022. doi:10.3390/computation10070127