

A novel multimodal hand database for biometric authentication

Bharath M. R.* and K. A. Radhakrishna Rao

Department of Electronics and Communication Engineering, PES College of Engineering, Mandya, Karnataka 571401, India

Received: 31-July-2021; Revised: 14-January-2022; Accepted: 15-January-2022

©2022 Bharath M. R. and K. A. Radhakrishna Rao. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Biometric authentication is one of the most exciting areas in the era of security. Biometric authentication ideally refers to the process of identifying or verifying the user through physiological and behavioral measurements using security processes. Multimodal biometrics are preferred over unimodal biometrics due to the defensive nature of multimodal biometrics. This research introduces a distinct hand database of individuals, which is acquired using a tailored hardware setup. The database contains four biometric traits: dorsal vein, wrist vein, palm vein, and palm print of the same person, which enables the multimodal biometric authentication exploration to create a spoof-proof authorization system. All four biometrics are captured using a single hardware device. The veins are highlighted by lighting up the infrared (IR) light emitting diodes (LEDs) and a complementary metal oxide semiconductor (CMOS) image sensor is used to capture the vein image. Similarly, the palm region is equally illuminated with the array of white LEDs, and the palm print is captured using the CMOS image sensor. The CMOS sensor has a simple structure and uses a single camera. A total of 308 participants were included in the database, resulting in 8336 unique hand vein and palm print images of both hands. Compared the captured database with the existing databases such as PUT, FYO, VERA, and Bosphorus in terms of availability of the traits, number of subjects, total number of images, number of sessions, and gender. Preliminary experiments were conducted on the self-acquired database using an open-source software package named "Orange". Various performance parameters were measured to construct a cost-effective authentication system. The classification accuracy obtained in k nearest neighbor (k NN), random forest (RF), support vector machine (SVM), neural network (NN), gradient boosting (GB), and logistic regression (LR) algorithms at a 70% learning rate are 99.8%, 99%, 99.8%, 99.8%, 97.7%, and 99.8%, respectively.

Keywords

Biometric authentication, Contrast limited adaptive histogram equalization, Complementary metal oxide semiconductor, Genuine, Hand vein, Imposter, Multimodal biometrics, Near infrared, Orange.

1. Introduction

Automated person identification through a biometric device is the most popular technology for individual authentication. Recognition is based on the behavioral or physical characteristics such as face, fingerprints, voice, stride, and signature [1]. These characteristics are known as biometric cues or biometric modalities. For the past several years, these biometrics cues have been used in applications ranging from less security to high security [2–5].

Due to the defensive nature of multimodal biometrics in terms of security and robustness, it is preferred over unimodal biometric systems [6].

Biometrics are simply a way to provide an assurance of authentication that a user cannot deny, and multibiometric improves authentication by combining information that is more unique to an individual than information received from a single source. The distinctive hand structure is due to the formation of bones, veins, muscles, and joints and the physiological correlation among them [7]. The internal and physiological anatomy of the human hand, which defines the individuality of several hand-associated biometric modalities, is the focus of this research. Palm print: The interior portion of the wrist to the starting point of the fingers forms the palmar region of a human hand. Palm print refers to the impression made by the palmar region when it is pressed against some surface. Palm prints mainly reveal features such as principal lines, wrinkle points, end points, width, length, and texture. The accuracy

*Author for correspondence

of palmprint identification is highly dependent on illumination, and spoofing is also possible. Hence, it is required to consider another biometric along with the palm print. Hence, in this research, hand veins are considered as another biometric to ensure accuracy, high user acceptance, and a spoof-proof system. Hand veins: human veins that reside under the skin carry deoxygenated blood from the tissues back to the heart. Veins are not easily identified by visual inspection systems. The blood vein patterns are so unique that even twins will have unique vein patterns [7]. Veins that are found in the wrist, palm and dorsal regions of the hand are of the precise interest in this research for biometric identification. Cephalic and Basilic are the two varieties of veins observed on the dorsal aspect of the hand. Bunch of veins that are attached to the wrist of the hand are referred to as cephalic veins, and veins that are attached to the surface on the dorsum side are referred to as basilic veins.

Vein patterns can be depicted through ridges and bifurcation points [8]. The physiological growth of a human impacts on the structure of the vascular network. These networks may shrink or extend over the ages. Thus, the shape of the hand veins updates accordingly. The vein patterns are steady between the ages of eighteen and fifty-five years due to the fact that the growth in adult life is almost null and the recognition of the vein patterns is easier than compared to earlier or later stages of the human life cycle. Initial vein patterns tend to change at later ages due to many factors, such as vessel shrinkage, degeneration of bones and muscles. The nature of diseases like diabetes, atherosclerosis, tumors, metabolic diseases, or hypertension can affect the human vascular system [9] and make it thick or thin. Recent trends in biometrics focus on human vein patterns and, in particular, hand vein patterns, which reside under the skin, which makes it difficult to spoof. The quality of the captured image plays a prime position in the success of hand vein recognition. The blood flowing through the vessels can be captured as dark pixels with the aid of a properly illuminated near infrared (NIR) source and the camera.

This research is motivated by the need to improve the accuracy of multimodal biometric systems in order to increase security. Furthermore, no database exists that has all 4 traits of unique individuals with age and gender information.

The objectives of this work are discussed below:

- Establish a novel database which includes 4 biometrics such as dorsal vein, wrist vein, palm vein, and palm print of both hands of an individual.
- Compare the acquired database with the existing database in terms of availability, number of subjects, total number of images, number of sessions, and gender.
- Performance evaluation of the self-acquired database.

The objectives are fulfilled by capturing the images through a customized image capturing hardware setup. Data acquisition was carried out in three sessions. The images have been gathered amongst volunteered college students and the employees/associates of People's Education Society College of Engineering (PESCE) and few were gathered from the public. Through the literature review, compared the self-acquired database with the existing database. Simulation experiments were carried out using these acquired biometric images and compared the results with the established classification algorithms.

The major contributions to this research work are:

- 1) Introduced one of the first multimodal biometric data, which contains a total of 8336 images from 308 subjects. The database includes dorsal vein, palm vein, wrist vein, and palm print images, as well as age and gender details of the same individuals.
- 2) Compared the self-acquired data with the existing publicly available database in terms of availability, number of subjects, total number of images, number of sessions, and gender.
- 3) Preliminary experiments were conducted using a new simulation tool on the self-acquired data. Various performance parameters were measured, which potentially reduced the software implementation cost.

The remaining sections are organized as below: section 2 provides a review of related literature, section 3 details the acquired hand database, section 4 provides the experimental assessment and discussion of the experiments is discussed in section 5, and section 6 provides the conclusion and recommendations for future work.

2.Literature review

Literature study mainly focuses on the types of available biometrics, databases, and their usage in biometric identification. Some parts of the literature

are focused on fusing biometrics and classification methods. A comprehensive comparison of biometric fusion is performed in [10]. Various testing methods were used to check the accuracy and speed of the system. In this study, the database of Institute of Automation Chinese Academy of Sciences (CASIA) v1.0 was utilized, which mainly contains multispectral palmprint images. Based on the 2-dimensional (2D) Gabor filter and a gradient-based approach, finger and palm vein biometric recognition was presented [11].

In 2020, authors introduced FYO database, which is a new multimodal database which is acquired in an administered environment using a medical vein finder. Applied feature extraction techniques such as Gabor filters, binarized statistical image features (BSIF), and histogram of oriented gradients (HOG). With two models, a deep learning-based convolutional neural network (CNN) architecture was presented, as well as fusion in the decision-level [12].

Various types of fusion schemes such as decision, score, sensor, and feature level were analyzed and compared. Authors also analyzed the performance metrics in [13].

Kabir et al. [14] presented a hybrid fusion technique at the feature level and new weighting called as mean extrema-based confidence weighting technique (MEBCW). This technique uses the scores which are generated from feature level fusion and lastly, they incorporated MEBCW in hash encoded features (HEF) to get the improved recognition rate. The claim rate of equal error rate (EER) and genuine acceptance rate (GAR) is 5.3% and 7.2% respectively.

A wrist identification algorithm was proposed and skin segmentation was performed, and feature extraction and classification were also included in the study. However, the authors showed their interest in localization of the key point and aligned the image to the template. In this study Nanyang Technological University (NTU) database version 1 is used, which contains the wrist images of 731 individuals and the total acquired image count was 3945[15].

Tiong et al. [16] published a new approach to multimodal face biometrics in 2020. This approach uses multi-feature fusion layers. It proved that the multi-feature fused layers increase accuracy as well as efficiency.

Mason et al. [17] proposed a biometric system for the healthcare environment. The CNN framework was used to achieve a true positive rate (TPR) of 98%. The authors compared the results to those of traditional feature extractors such as local binary patterns (LBP), HOG, and scale invariant feature transform (SIFT). Multiclass deep learning has been used in research on vein biometrics.

To identify the finger vein, a CNN based deep-learning approach is utilized in [18]. CNN based feature learning and transfer method has been applied to the hand dorsal vein [19]. As the deep learning approach required huge data to train the model, CNN based palm vein authentication was employed with pixelwise visibility-aware multi-view stereo network (PVSNet) architecture to train the model [20].

A multimodal biometric system is formed by the combination of finger shape and finger veins, and a CNN is incorporated into the work done by [21]. Features were extracted from the finger and finger veins using SIFT and deep CNN respectively to form a fusion based multi biometric system [22].

Abdalrady and Aly [23] used the sample face images from Gallagher's database for gender classification. In addition, they proposed that principal component analysis (PCANet) be used instead of CNN. Performance was compared with the state-of-the-art approaches. Deep CNN was used in the study for age and gender prediction of unfiltered faces.

In this study, authors addressed the variations of real-world unfiltered faces by incorporating robust image processing algorithms [24]. A comprehensive review of scientific literature, identifying the open challenges and offering suggestions for future research directions in soft biometrics, was provided by the authors in [25]. Age estimation of a person using compact multifeatured learning and fusion was proposed. A deep learning method has been proposed to overcome the limitations of the LBP, HOG, and BSIF methods. This study compared the age distributions of the MORPH2 face dataset with those of the FG-Net and looking at people (LAP) databases [26].

A detailed analysis of biometrics has been carried out in [27]. Additionally, author explained the different types of biometric devices and their applications and provided verification and identification in biometric methodology. Several biometric measures such as

pulse response, hand geometry, and finger vein biometrics [28].

On NIR images, a fuzzy rule-based system was used and obtained 92% accuracy. A new NIR hand image dataset, with 111,000 images overall, was presented [29]. It outlines an approach for enhancing the internal and external features of a finger. The finger region is analyzed using a distance transformation method, which allows the extraction of specific features. The method in this study was compared with widely used processing algorithms and achieved 90% accuracy. The 3σ criterion dynamic threshold strategy was used to localize the region of interest (ROI) in the finger vein in [30].

HKPU, MMBNU_6000, FV-USM, and ZSC-FV are the finger vein datasets they used, of which ZSC-FV was acquired locally and contains 37,080 finger vein samples. A Fusion of iris, face, and finger vein are captured in [31]. The SDUMLA-HMT database was utilized to develop the deep learning approach for the multimodal biometric recognition system and obtained an accuracy of 99.39%.

Lv et al. [32] used incompletely closed NIR (ICNIR) finger capturing equipment based on NIR imaging technique to perform feature-level fusion of finger vein and fingerprint.

Sarangi et al. [33] used ear and profile face in the improved multimodal biometric system. Histogram based local descriptors are used and fused them to get the high dimensional feature vector. The effectiveness of their proposed model was compared with the AlexNet and GoogleNet CNN models, and claimed that their proposed approach achieves better performance.

Vyas et al. [34] proposed accurate feature extraction techniques for iris and palm print biometrics. They proposed a coding-based approach named "bit transition code". They utilised the PolyU palm print database and the Indian Institute of Technology-Delhi (IITD) iris database. The comparison was made with the existing approaches and showcased the usefulness of their proposed work. Different fusion techniques, such as sensor level, feature level, score level, decision level, and hybrid fusion, are discussed, as well as their individual benefits and drawbacks [35].

The author demonstrated that the accuracy of the biometric system improves by using the multiple

features of a person. In this study, features from the iris, palm print, face, and finger veins were explored and obtained the good accuracy. In this study, the proposed multimodal CNN approach achieved an accuracy of 94% [36].

Varshini and Aravinth [37] proposed a hybrid fusion scheme that contains the electrocardiogram (ECG), face, and fingerprint modalities. The datasets used were FVC2002/2004, Face94, and PhysioNet. This fusion scheme obtained true positive rate (TPR), false positive rate (FPR), and equal error rate (EER) rates of 0.99, 0.1, and 0.5, respectively.

Leghari et al. [38] discussed the fusion of fingerprint and online signature. The CNN model was developed for early and late feature fusion, and in the early feature fusion obtained an accuracy of 99.1%, and in the late feature fusion obtained an accuracy of 98.35%.

Ghouzali et al. [39] used chaotic maps in multimodal biometric systems to address the cancelability issue. Face and fingerprint biometrics were utilized. Cancellable biometrics features were generated using logistic maps and torus automorphism. With this approach, they achieved a genuine acceptance rate of 100%.

However, it is evident from the aforementioned studies that there is a need for biometric combinations of the same individuals that have a high recognition rate and are spoof-proof. This research is motivated by the fact that spoofing of vein image is impossible since vein images can only be acquired from a living body, and the combination of vein images with the traditional hand surface images increases the security. Four biometrics of the same individuals, including palm vein, wrist vein, dorsal vein, and palm print, were established using the customised hardware configuration. Further sections go through the specifics of the self-acquired database.

3. Methods

3.1 Image acquisition

The capturing device is placed inside the bottom of the enclosure made from sunboard material. The overall height and width of the enclosure is 29cm \times 15cm respectively. An open window region of 8.5 cm \times 8.5 cm is made on the top of the enclosure to place the hand region. This window of the target (hand) is about 9-10 cm away from the camera source in such a way that a clear image is captured i.e., the focal length is adjusted. The inner portion of the enclosure

is made dark so that the illumination is proper throughout the image acquisition stage. The major advantage of this enclosure is that it has only two pegs at the top to guide the hand region to be aligned with the capturing device. Due to these pegs, hand rotation was minimized. To capture the dorsal vein, a notch is grooved on top of the enclosure near the pegs. The user must keep the dorsal side proximal phalange bone in the groove to avoid hand variations. The capturing device in this work contains four parts: 1) A CMOS camera. 2) White illumination LED for homogeneous illumination to capture palm print images. 3) NIR printed circuit board (PCB) (which includes NIR LEDs). 4) Small computer. A total of forty-eight NIR LEDs are arranged in a square form, with the camera at the center. This arrangement is used to acquire the palm, dorsal, and wrist veins. Six white LEDs are arranged in a strip to highlight the palm region and provide homogeneous illumination during the palm print capture. Processing and database storage are handled by a Raspberry®-based tiny computer. The Raspberry® Pi board has a 64-bit quad-core processor with a speed of 1.4GHz, dual-band wireless local area network (LAN) (2.4GHz and 5GHz), power over ethernet (PoE), bluetooth 4.2, and faster ethernet. This capturing device is connected via

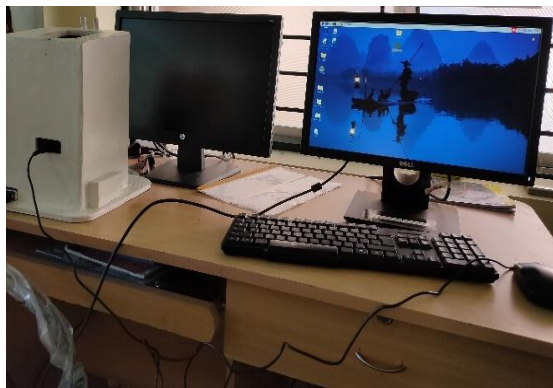


Figure 1 Image acquisition hardware setup

Environmental aspects:

The images are captured under unrestricted environmental conditions:

Temperature: Varied between 27–34 degree Celsius.

Humidity: Dry.

External light: Usually daytime, indoor only, and external artificial light (without direct sunlight).

Figure 3, *Figure 4* and *Figure 5* shows the acquisition procedures for palm vein and palm print, wrist vein, and dorsal vein, respectively. An image acquisition drive was conducted for 2 months in 3 different locations. A total of 308 volunteers stepped

video graphics array (VGA) connector to a 720p high definition (HD) monitor. The acquired images are displayed on the monitor and saved to the computer in joint photographic experts group (JPEG) format with a resolution of 1223×961 . Python 3.7.2 is used, which has the advantage of being rich in deep learning libraries. *Figure 1* shows the image acquisition hardware setup, and *Figure 2* shows the NIR LED arrangements with the CMOS camera at the center.

To acquire the image, the user is asked to perform the following steps:

- Insert the palmar side of the middle finger between the two pegs until the finger valley touches the pegs- to capture palm print and palm veins.
- Fold the forehand and keep the dorsal side proximal phalange bone in the groove -to capture dorsal vein.
- Keep the start of the wrist with the support of two pegs-to capture wrist veins.

All the above steps were followed for both hands.

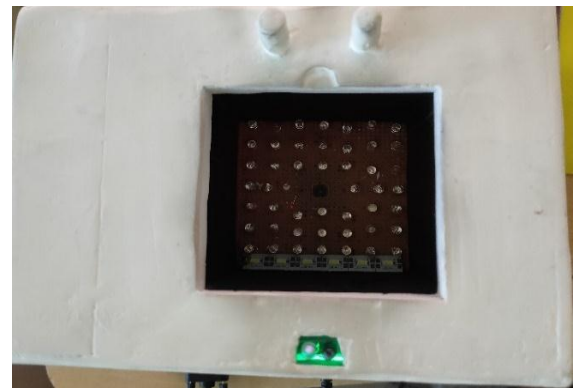


Figure 2 NIR LED arrangements with CMOS camera

forward in the image acquisition drive. Volunteers are college students, employees/associates of PESCE, and the general public. For each individual, the dorsal vein, wrist vein, palm vein, and palm print of both hands were acquired. An overall image count of 8336 was formed by capturing the biometric images of both hands. For a few individuals, three instances of each biometric were captured, while five instances were captured for others. The volunteers are in the age range of 15-76. As shown in *Table 1*, the volunteers ages are grouped into four categories, with males and females counted separately. A total of 161 male and 147 female datasets from different age

groups were acquired. *Figure 6* and *Figure 7* depict the graphical representation of the age distribution and age group representation, respectively (Total number=308). 48 percent of the collected data are between the ages of 20 and 30, and 32% are between the ages of 15-20. Similarly, 9 percent of the datasets are between the ages of 50 and 80, while 11 percent of the datasets are between the ages of 30 and 50. *Table 2* compares the collected database with the publicly available datasets such as PUT, FYO, VERA, Bosphorus. The number of subjects, male to female ratio, and samples per subject are presented.



Figure 3 Palm vein and palm print acquisition



Figure 4 Wrist vein acquisition

Each dataset includes the number of subjects, sessions, and samples per subject, as well as the male to female ratio. From the *Table 2*, it is concluded that the self-acquired data has a greater number of samples for experimentation as compared to other listed databases. Each person's data is stored in 308 individual folders, each with 4 sub folders, and each subfolder contains 2 more subfolders: one for left-hand data and another for right-hand data. All the captured images are stored in JPEG format.



Figure 5 Dorsal vein acquisition

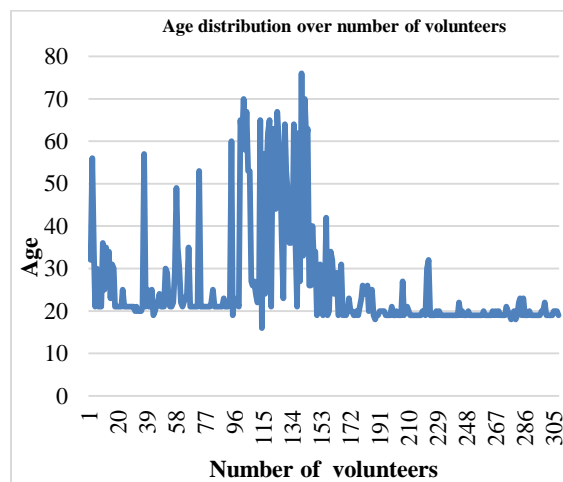


Figure 6 Age distribution over number of volunteers

The generic steps involved in image processing are as follows:

- Image capture.
- Pre-processing.
- Feature extraction.
- Matching.
- Determination/Decision.

Table 1 Age groups and gender distribution

Age group (Yrs.)	Males	Females	Total
15-20	44	54	98
20-30	80	67	147
30-50	20	14	34
50-80	17	12	29
Total	161	147	308

Figure 8 depicts the fundamental image processing block diagram. The dorsal vein, palm vein, wrist vein, and palm print of an individual were acquired using the customised hardware. Once the image is acquired, it is fed to the pre-processing stage where the image appearance is enhanced.

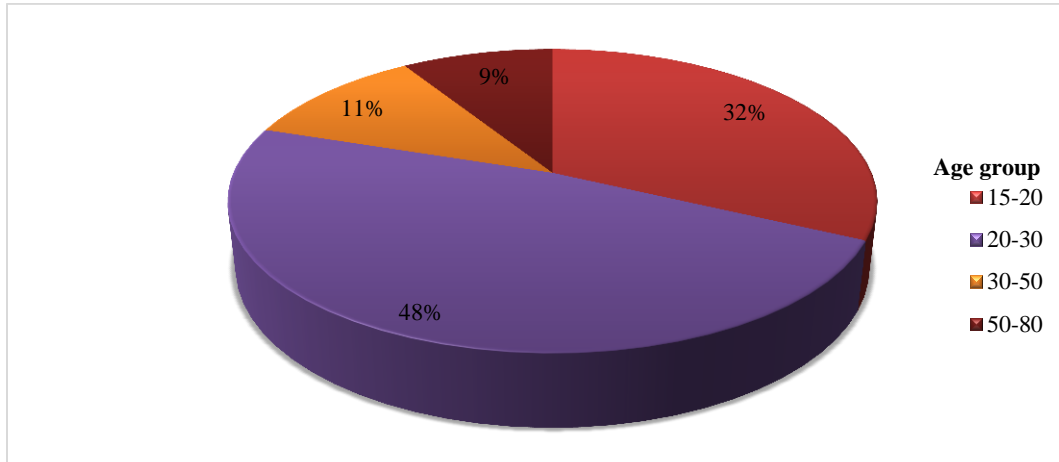


Figure 7 Age group in percentage

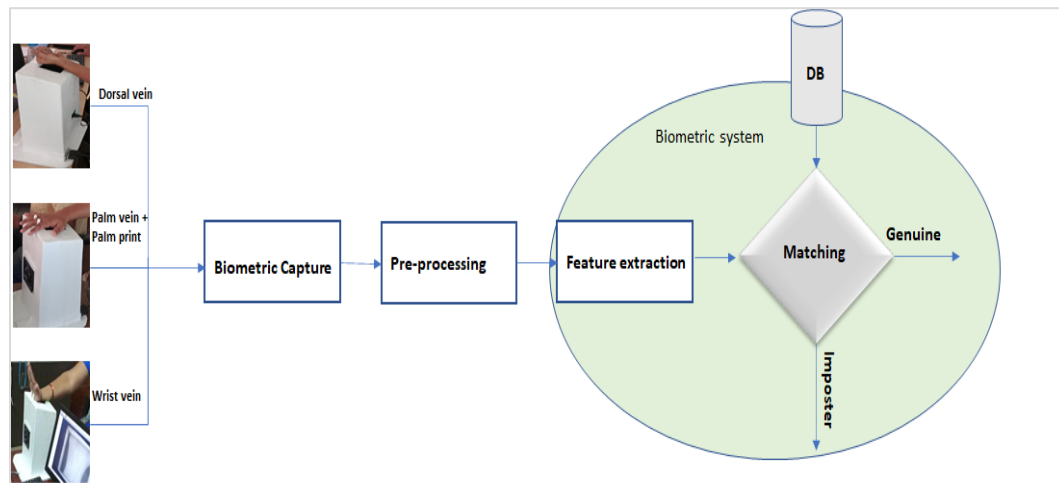


Figure 8 Image processing block diagram

Table 2 Database comparison (√ : Available, × : Not Available)

Database	PUT	FYO	VERA	Bosphorus	Self-collected
Palm vein	√	√	√	√	√
Dorsal vein	×	√	×	√	√
Wrist vein	√	√	×	×	√
Palm print	×	×	×	×	√
Subjects	50	160	110	100	308
Male/Female	×	111/49	70/40	58/42	161/147
Left/Right hands	2	2	2	2	2
Sessions	3	2	2	4	3
Total images	2400	1920	2200	1575	8336

3.2 Pre-processing

The main objective of pre-processing is to enrich and normalize the images to extract the unique features from it. The contrast limited adaptive histogram equalization (CLAHE) method is applied to increase the contrast between the veins and the rest of the hand region. The ROI size of 500×540 is used to extract the palmar, dorsal, and wrist regions, and the median filter is used to reduce the high frequency noise.

Following are the preprocessing procedure:

1. Read the gray scale image.
2. Extract the ROI.
3. Apply CLAHE.
4. Noise removal using median filter.

3.3 Contrast limited adaptive histogram equalization (CLAHE)

CLAHE was developed to improve the contrast of low-contrast medical images. The contrast limiting of CLAHE differs from that of standard adaptive histogram equalization (AHE). To combat the problem of noise amplification, the CLAHE added a clipping limit. Before computing the cumulative distribution function, the CLAHE restricts the amplification by histogram clipping at a predetermined value. In this method, histogram equalization is applied to all the contextual regions. The pixels that are clipped in the original histogram are reallocated to each gray level. Because each pixel intensity is confined to a specified maximum, the redistributed histogram differs from the generic histogram. The improved and original images have the same minimum and maximum gray values. Block size (BS) and clip limit (CL) are the two important parameters to be considered while applying CLAHE.

Algorithm steps:

- Step 1: Obtain non-overlapping contextual sections by dividing the original intensity images.
- Step 2: Calculate each contextual section's histogram based on the gray levels in the array image.
- Step 3: Calculate the contrast limited histogram of the contextual section by the CL value using the below equations:

Average number of pixels is calculated by Equation 1.

$$I_{Avg} = (I_X \times I_Y) / I_{Gray} \quad (1)$$

Where,

I_{Gray} defines the number of gray levels in the contextual section.

I_X = The number of pixels in the contextual section's X dimension.

I_Y = The number of pixels in the contextual section's Y dimension.

Clip limit is given by the Equation 2.

$$I_{CL} = I_{clip} \times I_{Avg} \quad (2)$$

Where, I_{CL} = Actual clip limit.

I_{Avg} = Average number of pixels.

I_{clip} = Normalized clip limit and its range are [0, 1].

If the actual clip limit I_{CL} is less than the number of pixels, then the pixels are clipped.

Equation 3 represents the average pixel values to distribute to each gray level.

$$I_{Avggray} = I_{\Sigma clip} / I_{Gray} \quad (3)$$

Where, $I_{\Sigma clip}$ = Sum of clipped pixels.

Step 4: Re-distribute the left-out pixels till all of them are distributed. Then, the redistributed pixel is given by Equation 4.

$$I_{Redistr} = I_{Gray} / I_{Remain} \quad (4)$$

Where, I_{Remain} = Remaining pixels

Step 5: Apply a linear transform to enrich the intensity values in each region.

Step 6: Apply the bilinear interpolation method between the 4 alternative mappings to calculate the redistributed pixel values and thereby remove the boundary artifacts.

3.4 Median filtering

Noises might occur throughout the image capture and transmission process. Gaussian noise, salt and pepper noise, and impulsive noise are the most common types of noise. The median filter is a commonly used nonlinear filter to remove the noise that occurs during the image capture process. The major advantage of median filtering is its edge-preserving capability. The median filter is a statistically based nonlinear signal processing method. The median value of the neighborhood replaces the noisy value of the digital image or sequence (mask). The mask's pixels are ranked in order of their gray levels, and the group's median value is saved to replace the noisy value. The median filtering output is given by Equation 5.

$$g(x, y) = med\{f(x - i, y - j) | i, j \in W\} \quad (5)$$

Where, $f(x, y)$ = Input image, $g(x, y)$ = Output image.

W is the 2-dimension mask of the size $n \times n$ in the shape that may be circular, linear, square, etc. The

commonly used mask sizes are odd numbers, such as 3×3 , 5×5 , etc. In rare cases, even order median filters are used. The output of an even-order median filter is the average of the middle two samples.

4.Results

Figure 9 shows the preprocessed images of dorsal vein, palm vein, wrist vein, palm print. All four traits are preprocessed independently. Imposter datasets are created by taking the laser color print of the original acquisition and feeding them to the same image capturing device and preprocessing them further. The original pre-processed image set is labelled as genuine, and the other pre-processed image set is labelled as an imposter. The preprocessing of biometric images is implemented in Python v3.7. To evaluate the acquired datasets, a python-based simulator tool named "Orange" version 3.28.0 is used. In this research, 100 genuine and 97 imposter data sets were utilized for experimentation (altogether, 3072 images). These preprocessed genuine and imposter images are given to the "Import Images" block in the "Orange" tool, followed by the image embedding block, where this block reads the image and uploads it to the remote server or evaluates it locally.

The image embedding is done through deep neural networks. Feature vectors are calculated for each

image through these deep learning models and return the image descriptor. In this study, the SqueezeNet embedder is utilized, which is a deep model for image recognition. This widget automatically retrieves the image vectors from the server.

The SqueezeNet embedder evaluates faster compared to other embedders. The input for this block is a list of images, and the output is images represented in a vector. Figure 10 depicts the image embedder widget. The obtained embedded image parameters are given to the rank widget, where the ranking of attributes for classification takes place. The input to this widget is embedded images, and the outputs are reduced data, scores, and attribute lists. This rank widget scores the variables in step with its correlation with discrete or numeric variables, primarily on the relevant internal or external scorers. The ReliefF feature selection approach is utilized since this study is interested in binary classification of the gathered dataset. ReliefF calculates a proxy statistic for every feature that may be used to estimate the characteristic quality. This characteristic information is known as characteristic weights or, more widely, known as characteristic or feature scores, which could vary from -1 to +1. In total, 21 best attributes were selected according to their rankings. Figure 11 shows the selection of scoring methods and attributes.

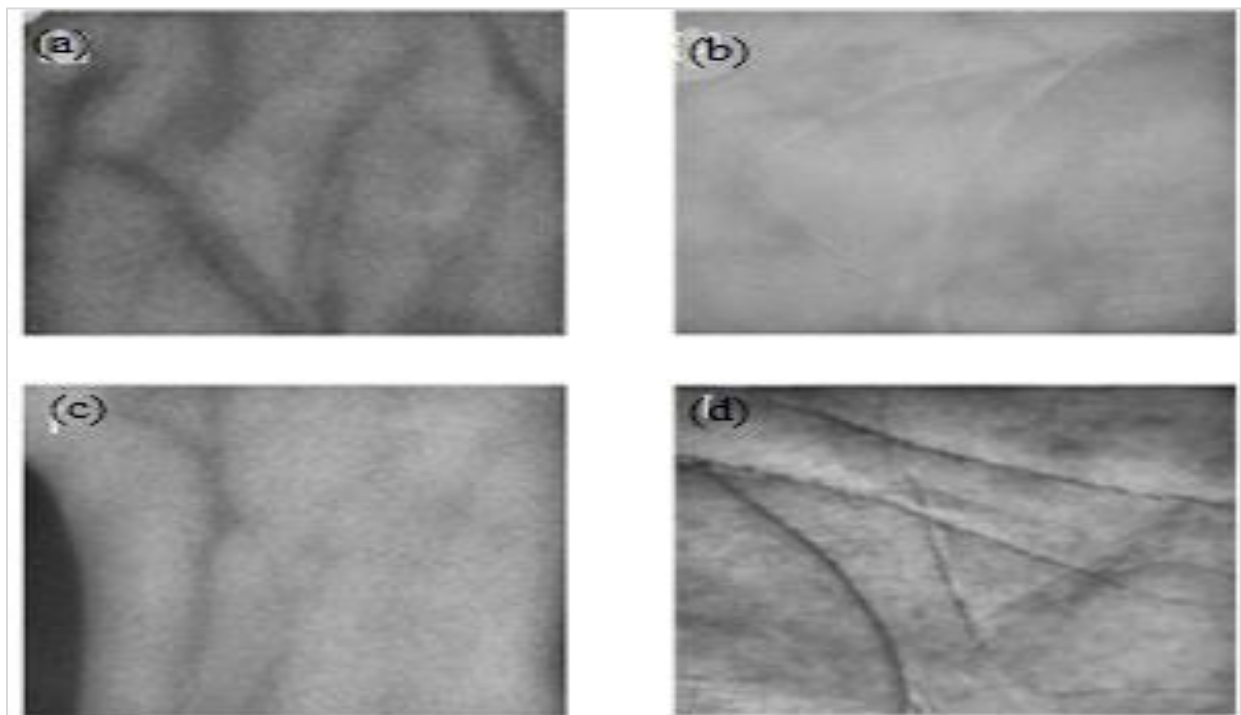


Figure 9 Pre-processed biometric images(a) Dorsal vein (b) Palm vein (c) Wrist vein (d) Palm print

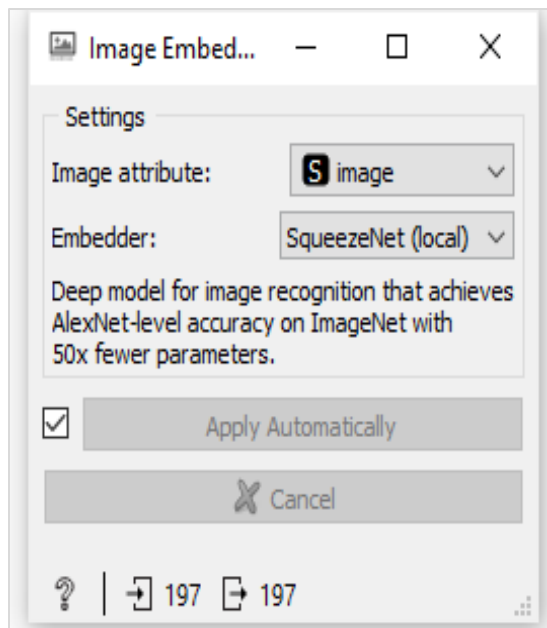


Figure 10 Widget to select image embedder

The selected attributes are passed to the test and score widget, where the widget tests the learning algorithms. In this work, compared the most commonly used learning algorithms, such as support vector machine (SVM), random forest (RF), k nearest neighbor (kNN), gradient boosting (GB), logistic regression (LR), and neural network (NN). The input to the test and score widget is a dataset and learning algorithms, and the outcome of this widget will be evaluated results. The radial basis function (RBF) kernel is used in the SVM method, with a numerical tolerance of 0.0010 and an iteration limit of 100. In the NN, neurons in the hidden layer are set to 100, rectified linear unit (ReLU) activation function and Adam optimization technique is applied, and the number of iterations is set to 100. In LR, the ridge(L2) regularization type is set and the strength value of 0.008 is set. In the kNN, Euclidian distance is considered as a metric and uniform weight is applied. In the GB method, number of trees is set to 100 with the growth control parameters such as depth of individual trees and subset split being set to 3 and 2 respectively.

Table 3, Table 4 and Table 5 depicts the overall performance of the selected classifiers at the different learning percentages. A random sampling method is used to validate each model. The data is separated into training and testing sets in a 60:40, 70:30, and 80:20 ratio, and the procedure is repeated ten times. The performance of each model is represented by the parameters F1 score, classification accuracy, 136

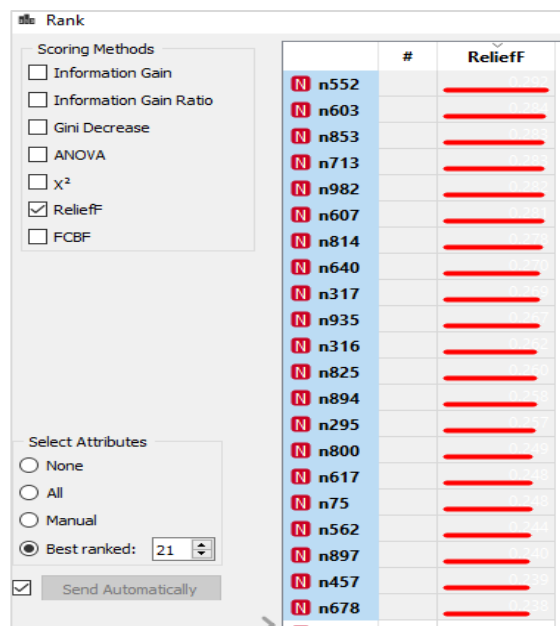


Figure 11 Scoring method and attribute selection

precision, recall, and specificity. The F1 score is the degree of weighted harmonic mean of precision and recall. Classification accuracy is the degree of effective classified samples. Precision is the ratio of authentic positives among instances categorized as positive. Recall is the ratio of authentic positives among all positive instances in the data. Specificity is the percentage of real negatives that had been anticipated as negative. The F1 score, classification accuracy, precision, recall, and specificity of the kNN, SVM, NN, and LR algorithms are 0.995 at a learning rate of 60. In terms of classification accuracy, these methods outperformed RF and GB by 1.3% and 3.8%, respectively. kNN, SVM, NN, and LR classifiers obtained 0.998 classification accuracy and F1 scores at a learning rate of 70. The precision and specificity of all the classifiers except GB show maximum values of 1.00. These methods achieved an improvement of 0.8% and 2.1% over the RF and GB methods in terms of classification accuracy. Similarly, at a learning rate of 80, kNN, SVM, NN, and LR classifiers achieved 0.995 classification accuracy and F1 scores. These methods achieved an improvement of 1.0% and 2.0% over the RF and GB methods in terms of classification accuracy.

Figure 12 and Figure 13 depict the receiver operating characteristics (ROC) curve of the selected prediction algorithms for genuine and imposter target classes, respectively. It serves as a mean of the comparison between classification models. It is the plot of the false positive (FP) rate on the x-axis versus the true

positive (TP) rate on the y-axis. In both the target classes, the ROC is nearer to the value of 1 for kNN, SVM, NN, and LR classifiers. Hence, these classifiers are expected to be more optimal. Classification accuracy comparison of the unimodal and multimodal biometric systems at a 70% learning rate is depicted in *Table 6*.

5. Discussion

The success of any biometric authentication system relies on the quality of the acquired images. To add to the discussion, the quality hand veins and palm print images are obtained from the capturing device for the following reasons:

- a) A well-controlled environment. The capturing device enclosure is designed in such a way that the effect of external lighting conditions is minimal on the captured hand images.
- b) Reduced hand rotation. A unique two-peg and groove structure minimized the hand rotation during the capturing process.

Table 7 depicts the performance of the FYO database at a learning rate of 70%. A total of 2248 images from the FYO database and 3072 images from the self-acquired database of genuine and imposter categories are considered in the simulation process. The comparison of *Table 4* and *Table 7* shows the supremacy of the self-acquired database over the FYO database in terms of different performance metrics such as F1 score, classification accuracy, training time, precision, recall, and specificity for different learning algorithms. In comparison to the FYO database, the better classification results were achieved in the self-acquired database due to the above-mentioned factors, i.e., the well-controlled environment and the reduced hand rotation. Based on all the experiments and comparisons, the kNN, SVM, NN, and LR methods achieved higher classification accuracy for self-acquired datasets at different learning rates. The time required to train the kNN, SVM, RF, NN, LR, and GB models with self-acquired datasets is 0.255, 0.436, 0.353, 4.239, 0.342, and 2.967 seconds, respectively, whereas the FYO database consumed 0.224, 0.619, 0.422, 16.243, 0.690, and 7.900 seconds for the same learning models. As it is seen in the comparative analysis of unimodal and multimodal in *Table 6*, the multimodal biometric combination achieved better classification accuracy compared to the unimodal biometrics in

different learning models. The dorsal vein achieved a classification accuracy of 99.5%, 99.5%, 98.5%, 99.5%, 99.5%, and 97.5% in kNN, SVM, RF, NN, LR, and GB methods, respectively. The palm vein achieved a classification accuracy of 97.8%, 97.8%, 97.8%, 99.5%, 99.5%, and 97.5% in kNN, SVM, RF, NN, LR, and GB methods, respectively. Similarly, wrist vein and palm print achieved a classification accuracy of 99.8%, 99.8%, 98.5%, 99.8%, 99.8%, and 98.5% in kNN, SVM, RF, NN, LR, and GB methods, respectively. Whereas the combination of dorsal vein, palm vein, wrist vein, and palm print achieved a classification accuracy of 99.8%, 99.8%, 99.0%, 99.8%, 99.8%, and 97.7% in kNN, SVM, RF, NN, LR, and GB methods, respectively.

The facts of the literature on multimodal biometrics availability and the better performance were achieved in this study. A unique combination of wrist vein, palm vein, dorsal vein, and the traditional palm print increases the performance of multimodal biometric recognition on the self-acquired database in most of the learning models, making this database independent of different learning algorithms. Overall, the literature review, experiments, and results obtained for the self-acquired database have demonstrated that it can be used in high-security biometric authentication requirements to defend against biometric attacks. A list of abbreviations is shown in *Appendix I*.

5.1 Limitations

Despite the success in terms of performance for the newly acquired dataset, a few limitations of the simulator could be considered. In this study, an open-source simulator named "Orange" is used, where it extracts the complete imaging attributes (1024) and the classification is performed based on the selected best ranked attributes. However, when dedicated features of each trait are extracted and processed, the overall classification accuracy is expected to vary with the large datasets.

Very few samples are available between the ages of 30-80 years in the self-acquired database. When age is used as a soft biometric, the training model will get fewer samples in this age group. Data augmentation techniques could be utilized to increase the number of samples in the age range of 30-80 years.

Table 3 Performance of self-acquired database at 60% learning rate

Model	F1	Classification accuracy	Training time(s)	Precision	Recall	Specificity
kNN	0.995	0.995	0.185	0.995	0.995	0.995
SVM	0.995	0.995	0.337	0.995	0.995	0.995
RF	0.982	0.982	0.245	0.983	0.982	0.983
NN	0.995	0.995	8.136	0.995	0.995	0.995
LR	0.995	0.995	0.561	0.995	0.995	0.995
GB	0.957	0.957	2.645	0.957	0.957	0.958

Table 4 Performance of self-acquired database at 70% learning rate

Model	F1	Classification accuracy	Training time(s)	Precision	Recall	Specificity
kNN	0.998	0.998	0.255	1.000	0.997	1.000
SVM	0.998	0.998	0.436	1.000	0.997	1.000
RF	0.990	0.990	0.353	1.000	0.981	1.000
NN	0.998	0.998	4.239	1.000	0.997	1.000
LR	0.998	0.998	0.342	1.000	0.997	1.000
GB	0.977	0.977	2.967	0.987	0.968	0.987

Table 5 Performance of self-acquired database at 80% learning rate

Model	F1	Classification accuracy	Training time(s)	Precision	Recall	Specificity
kNN	0.995	0.995	0.203	0.995	0.995	0.995
SVM	0.995	0.995	0.364	0.995	0.995	0.995
RF	0.985	0.985	0.247	0.985	0.985	0.985
NN	0.995	0.995	9.618	0.995	0.995	0.995
LR	0.995	0.995	0.563	0.995	0.995	0.995
GB	0.975	0.975	3.300	0.975	0.975	0.975

Table 6 Classification accuracy of unimodal and multimodal system at 70% learning rate

Model	Dorsal vein classification accuracy	Palm vein classification accuracy	Wrist vein classification accuracy	Palm print classification accuracy	DV+PV+WV+PP (classification accuracy)
kNN	0.995	0.978	0.998	0.998	0.998
SVM	0.995	0.978	0.998	0.998	0.998
RF	0.985	0.978	0.985	0.985	0.990
NN	0.995	0.995	0.998	0.998	0.998
LR	0.995	0.995	0.998	0.998	0.998
GB	0.975	0.975	0.985	0.985	0.977

Table 7 Performance of FYO database at 70 % learning rate

Model	F1	Classification accuracy	Training time(s)	Precision	Recall	Specificity
kNN	0.997	0.997	0.224	0.997	0.997	0.998
SVM	0.997	0.997	0.619	0.997	0.997	0.997
RF	0.995	0.995	0.422	0.995	0.995	0.995
NN	0.996	0.996	16.243	0.996	0.996	0.996
LR	0.996	0.996	0.690	0.996	0.996	0.997
GB	0.995	0.995	7.900	0.995	0.995	0.995

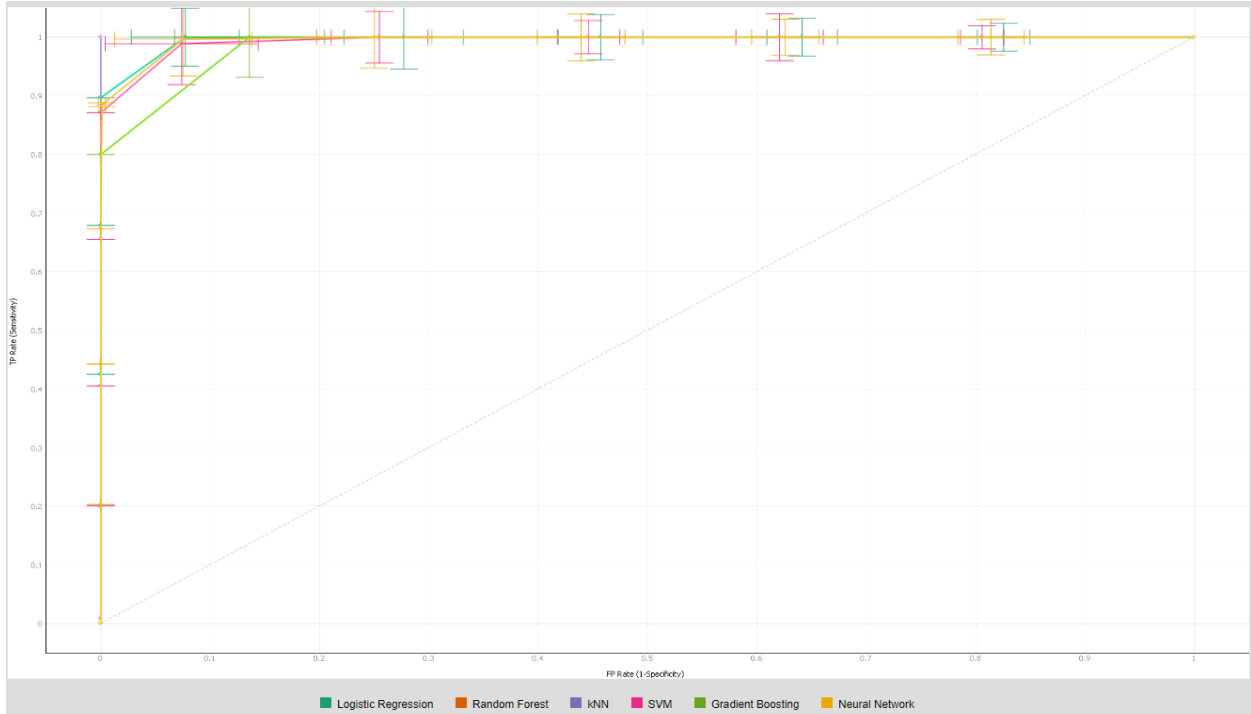


Figure 12 ROC analysis for genuine target class

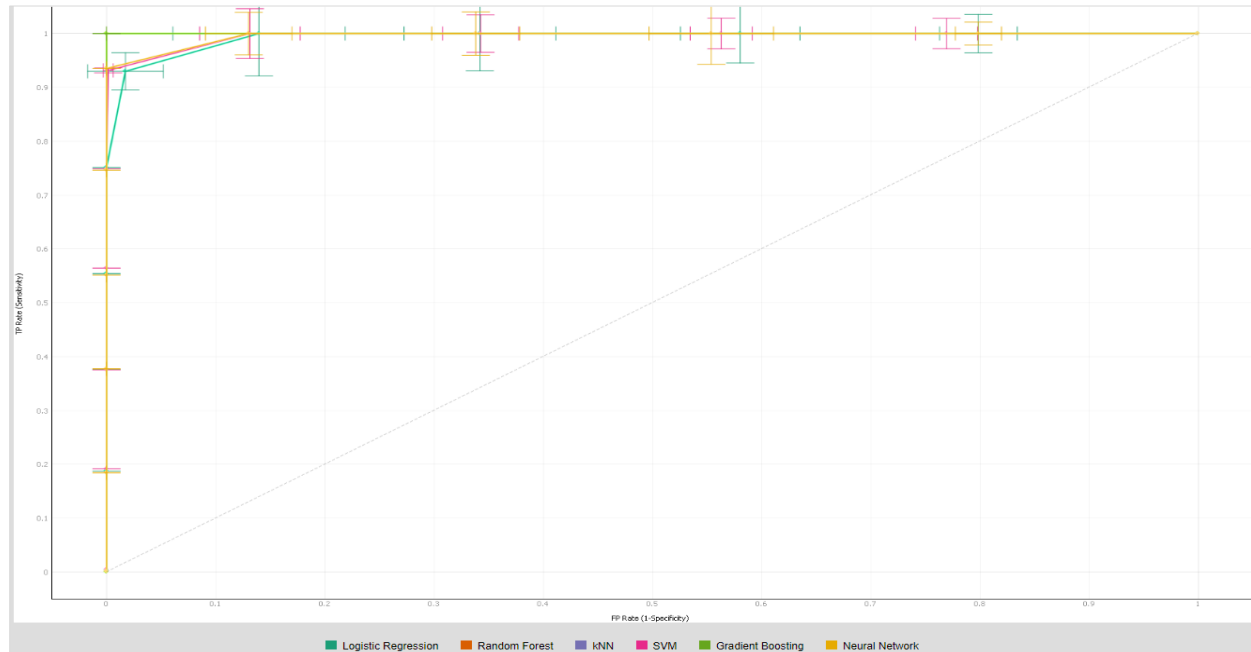


Figure 13 ROC analysis for imposter target class

6. Conclusion and future work

This research introduced a new multimodal hand database that includes an individual's palm vein, wrist vein, dorsal vein, and palmprint. This unique combination of biometric dataset was acquired

locally using a single customised NIR based portable capturing hardware unit. Having this self-acquired database of the same individuals, multimodal biometric research areas can be explored. Further, compared the self-acquired database with the existing

publicly available databases such as PUT, FYO, VERA, and Bosphorus in terms of availability, number of subjects, gender, number of sessions, and total number of images. Preliminary experiments were carried out using the self-acquired database. The performance parameters such as F1 score, classification accuracy, precision, recall, and specificity were measured. Experiment results of the unimodal and multimodal combinations (dorsal vein, palm vein, wrist vein, and palm print) showed that the proposed multimodal combination achieved the best result compared to the individual biometric traits. The well-controlled environment set up and the reduced hand rotation during the image acquisition enhanced the performance of the self-acquired database. The classification algorithms such as kNN, SVM, RF, GB, NN, and LR were experimented. The simulation outcome in terms of classification accuracy and F1 score of 99.8% was achieved in kNN, SVM, NN, and LR algorithms. Similarly, RF and GB algorithms achieved a decent classification accuracy and F1 score value of 99% and 97.7%, respectively, and boosted the confidence in implementing them in real time. Overall, the experiments have reduced software development and implementation costs. The self-acquired database has been proven to be the most suitable for biometric image classification. The self-acquired database also has other soft biometric details such as age and gender, hence, these details will be explored in future studies. The dedicated features from each biometric will be extracted, and the simulated learning algorithms will be used for biometric authentication. Also, deep learning models will be investigated for different security levels and different biometric combinations.

Acknowledgment

Authors would like to acknowledge Dhruva Kumar L, scholar and Ö. Toygar for their help and the volunteers who extended their support for data acquisition process during the pandemic situation.

Conflicts of interest

The authors have no conflicts of interest to declare.

Authors contribution statement

Bharath M.R.: Data collection, conceptualization, design, investigation on challenges, writing – original manuscript draft, analysis and interpretation of results. **Dr. K. A. Radhakrishna Rao:** Study conception, design, supervision, and review of draft manuscript.

References

- [1] Jain AK, Ross AA, Nandakumar K. Introduction to biometrics. Springer Science & Business Media; 2011.
- [2] Jain AK, Li SZ. Handbook of face recognition. New York: Springer; 2011.
- [3] Bowyer KW, Burge MJ. Handbook of iris recognition. Springer London; 2016.
- [4] Maltoni D, Maio D, Jain AK, Prabhakar S. Handbook of fingerprint recognition. Springer Science & Business Media; 2009.
- [5] Jain AK, Nandakumar K, Ross A. 50 years of biometric research: accomplishments, challenges, and opportunities. *Pattern Recognition Letters*. 2016; 79:80-105.
- [6] Dahea W, Fadewar HS. Multimodal biometric system: a review. *International Journal of Research in Advanced Engineering and Technology*. 2018; 4(1):25-31.
- [7] Li SZ, Jain A. Encyclopedia of biometrics. Springer Publishing Company, Incorporated; 2015.
- [8] Kumar A, Prathyusha KV. Personal authentication using hand vein triangulation. In *biometric technology for human identification V 2008*. International Society for Optics and Photonics.
- [9] Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. *Nature*. 2000; 407(6801):249-57.
- [10] Singh M, Singh R, Ross A. A comprehensive overview of biometric fusion. *Information Fusion*. 2019; 52:187-205.
- [11] Bharathi S, Sudhakar R. Biometric recognition using finger and palm vein images. *Soft Computing*. 2019; 23(6):1843-55.
- [12] Toygar Ö, Babalola FO, Bitirim Y. FYO: a novel multimodal vein database with palmar, dorsal and wrist biometrics. *IEEE Access*. 2020; 8:82461-70.
- [13] Verma SB, Saravanan C. Performance analysis of various fusion methods in multimodal biometric. In *international conference on computational and characterization techniques in engineering & sciences 2018* (pp. 5-8). IEEE.
- [14] Kabir W, Ahmad MO, Swamy MN. Weighted hybrid fusion for multimodal biometric recognition system. In *international symposium on circuits and systems 2018* (pp. 1-4). IEEE.
- [15] Matkowski WM, Chan FK, Kong AW. A study on wrist identification for forensic investigation. *Image and Vision Computing*. 2019; 88:96-112.
- [16] Tiong LC, Kim ST, Ro YM. Multimodal facial biometrics recognition: dual-stream convolutional neural networks with multi-feature fusion layers. *Image and Vision Computing*. 2020.
- [17] Mason J, Dave R, Chatterjee P, Graham-allen I, Esterline A, Roy K. An investigation of biometric authentication in the healthcare environment. *Array*. 2020.
- [18] Das R, Piciuccio E, Maiorana E, Campisi P. Convolutional neural network for finger-vein-based biometric identification. *IEEE Transactions on Information Forensics and Security*. 2018; 14(2):360-73.

- [19] Al-johania NA, Elrefaei LA. Dorsal hand vein recognition by convolutional neural networks: feature learning and transfer learning approaches. *International Journal of Intelligent Engineering and Systems*. 2019; 12(3):178-91.
- [20] Thapar D, Jaswal G, Nigam A, Kanhangad V. PVSNet: palm vein authentication siamese network trained using triplet loss and adaptive hard mining by learning enforced domain specific features. In *international conference on identity, security, and behavior analysis 2019* (pp. 1-8). IEEE.
- [21] Kim W, Song JM, Park KR. Multimodal biometric recognition based on convolutional neural network by the fusion of finger-vein and finger shape using near-infrared (NIR) camera sensor. *Sensors*. 2018; 18(7):1-34.
- [22] Kovac I, Marák P. Multimodal biometric system based on fingerprint and finger vein pattern. *Slovak Univ. Technol., Bratislava, Slovakia*. 2019.
- [23] Abdalrady NA, Aly S. Fusion of multiple simple convolutional neural networks for gender classification. In *international conference on innovative trends in communication and computer engineering 2020* (pp. 251-6). IEEE.
- [24] Agbo-ajala O, Viriri S. Deeply learned classifiers for age and gender predictions of unfiltered faces. *The Scientific World Journal*. 2020.
- [25] Hassan B, Izquierdo E, Piatrik T. Soft biometrics: a survey. *Multimedia Tools and Applications*. 2021:1-44.
- [26] Deng Y, Teng S, Fei L, Zhang W, Rida I. A multifeature learning and fusion network for facial age estimation. *Sensors*. 2021; 21(13):1-15.
- [27] Pujari MV, Patil R, Sutar MS. Research paper on biometrics security. *Contemporary Research in India*. 2021.
- [28] Haider SA, Rehman Y, Ali SM. Enhanced multimodal biometric recognition based upon intrinsic hand biometrics. *Electronics*. 2020; 9(11):1-20.
- [29] Bernacki K, Morón T, Popowicz A. Modified distance transformation for image enhancement in NIR imaging of finger vein system. *Sensors*. 2020; 20(6):1-9.
- [30] Yao Q, Song D, Xu X. Robust finger-vein ROI localization based on the 3σ criterion dynamic threshold strategy. *Sensors*. 2020; 20(14):1-21.
- [31] Alay N, Al-baity HH. Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits. *Sensors*. 2020; 20(19):1-17.
- [32] Lv GL, Shen L, Yao YD, Wang HX, Zhao GD. Feature-level fusion of finger vein and fingerprint based on a single finger image: the use of incompletely closed near-infrared equipment. *Symmetry*. 2020; 12(5):1-18.
- [33] Sarangi PP, Nayak DR, Panda M, Majhi B. A feature-level fusion based improved multimodal biometric recognition system using ear and profile face. *Journal of Ambient Intelligence and Humanized Computing*. 2021:1-32.
- [34] Vyas R, Kanumuri T, Sheoran G, Dubey P. Accurate feature extraction for multimodal biometrics combining iris and palmprint. *Journal of Ambient Intelligence and Humanized Computing*. 2021:1-9.
- [35] Bala N, Gupta R, Kumar A. Multimodal biometric system based on fusion techniques: a review. *Information Security Journal: A Global Perspective*. 2021:1-49.
- [36] Vijayakumar T. Synthesis of palm print in feature fusion techniques for multimodal biometric recognition system online signature. *Journal of Innovative Image Processing*. 2021; 3(2):131-43.
- [37] Varshini SA, Aravinth J. Hybrid level fusion schemes for multimodal biometric authentication system based on matcher performance. In *computational vision and bio-inspired computing 2021* (pp. 431-47). Springer, Singapore.
- [38] Leghari M, Memon S, Dhomeja LD, Jalbani AH, Chandio AA. Deep feature fusion of fingerprint and online signature for multimodal biometrics. *Computers*. 2021; 10(2):1-15.
- [39] Ghouzali S, Nafea O, Wadood A, Hussain M. Cancelable multimodal biometrics based on chaotic maps. *Applied Sciences*. 2021; 11(18):1-15.



Bharath M. R. received his B.E. and M.Tech degrees in the stream of Electronics and Communication Engineering from Visvesvaraya Technological University, Belagavi, Karnataka India, in 2007 and 2010 respectively. Current research interests are in the area of Biometrics, Computer

Vision, and Image Processing.
Email: bharathmullur@gmail.com



Dr. K. A. Radhakrishna Rao obtained his doctoral degree from Indian Institute of Science, Bangalore in the area of Biomedical Signal processing and Master's from Indian Institute of Technology, Madras with specialization in IC circuits and Systems. Dr. Rao is associated with P E

S College of Engineering for more than 30 years right from his Bachelor's degree and currently he is professor in the department of Electronics and Communication Engineering. He has a research experience of more than 20 years and has 52 publications to his credit. Dr. Rao has large interest in Signal Processing and Biometrics and currently guiding four students for doctoral degree. He was part of many international conferences, review member for many journals and expert member in panel discussion.
Email: karkrao@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	AHE	Adaptive Histogram Equalization
2	BS	Block Size
3	BSIF	Binarized Statistical Image Features
4	CASIA	Chinese Academy of Sciences Institute of Automation
5	CL	Clip Limit
6	CLAHE	Clip Limit Adaptive Histogram Equalization
7	CMOS	Complementary Metal Oxide Semiconductor
8	CNN	Convolutional Neural Networks
9	ECG	Electrocardiogram
10	EER	Equality Error Rate
11	FP	False Positive
12	FPR	False Positive Rate
13	GAR	Genuine Acceptance Rate
14	GB	Gradient Boosting
15	HEF	Hash Encoded Features
16	HD	High Definition
17	HoG	Histogram of Oriented Gradients
18	ICNIR	Incompletely Closed Near Infra-Red
19	IITD	Indian Institute of Technology-Delhi
20	JPEG	Joint Photographic Experts Group
21	kNN	K-Nearest Neighbour
22	LAN	Local Area Network
23	LAP	Looking At People
24	LBP	Local Binary Patterns
25	LED	Light Emitting Diode
26	LR	Logistic Regression
27	MEBCW	Mean Extrema Based Confidence Weighting
28	NIR	Near Infra-Red
29	NTU	Nanyang Technological University
30	NN	Neural Network
31	PCANet	Principal Component Analysis Network
32	PCB	Printed Circuit Board
33	PESCE	People's Education Society College of Engineering
34	PoE	Power over Ethernet
35	PVSNet	Pixelwise Visibility aware multi-view Stereo Network
36	RBF	Radial Basis Function
37	ReLU	Rectified Linear Unit
38	RF	Random Forest
39	ROC	Receiver Operating Characteristics
40	ROI	Region of Interest
41	SIFT	Scale Invariant Feature Transform
42	SVM	Support Vector Machine
43	TP	True Positive
44	TPR	True Positive Rate
45	VGA	Video Graphics Array
46	2D	2 Dimension