Academic Science Journal

Evaluation of the Proposed Hand Vein Authentication System Using Machine Learning

Rajaa Ahmed Ali [1*] and Ziyad Tariq Mustafa Al-Ta'i [2]

[1]Department of Computer Science, The Institute of informatics for post-graduation, University of Information Technology and Communication

[2]Department of Computer Science, College of Science, University of Diyala

[*]Phd202120689@iips.edu.iq

Received: 6 December 2024 **Accepted: 7 January 2025**

Published: 28 April 2025

DOI: https://dx.doi.org/10.24237/ASJ.03.02.960B

Abstract

This paper discusses the use of a biometric system for secure authentication based on deeper exploration in hand vein patterns to derive more reliability as a good biometric trait. The proposed system identifies internal hand vein patterns, and also identifies individuals correctly throughout their lifetimes. The proposed system follows a three-phase approach comprising data processing, feature extraction, and feature evaluation. For the advanced filtering techniques in a dataset of 420 hand vein images, the approach goes to Box filters, HFEF, and CLAHE, then to discrete cosine transformation and image fusion. The features are extracted through PCA Net for acquiring the most distinctive attributes of hand veins. The different machine learning algorithms used in this evaluation for classification of the extracted features include SVM, Logistic Regression, and Naive Bayes. Results indicate the highest accuracy for the Logistic Regression algorithm (99.7%) and the SVM algorithm of (99.6%). However, the Random Forest algorithm has an accuracy of (98%), while the Naive Bayes algorithm shows a poorer accuracy of (91%).

Keywords: Biometric authentication, hand vein patterns, machine learning, feature extraction.

Academic Science Journal

Introduction

Biometric systems, on the other hand, define or authenticate individuals by using their biological attributes or their behavioral characteristics measured against biometric designs in the database of similar attributes. Hence, this measurement can be classified as an identification or verification system and can be defined as a computerized method of identifying one's identity or verifying it through specific biological or behavioral characteristics [1]. These are the two unique kinds of biometric characteristics that do not vary over time: physiological and behavioral. The physiological features include the fingerprint, DNA, iris, hand, face, etc., while behavior includes voice, signature, and keystroke, etc., as shown in Figure 1.[2].

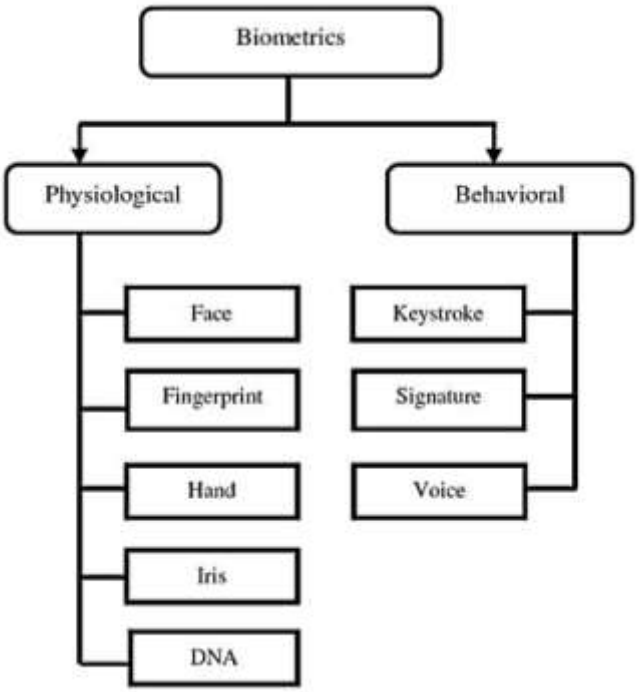

Figure 1 : Biometric systems general block diagram [2].

The fact is that the palm vein pattern be used as a biometric trait for authentication as this pattern has spatial geometry of variable dimensions for each user which cannot be amended [3].

The interior part of the hand behind the thumb and index finger is called the palm. The image of the correct hand held on the flat, glass surface of a scanner is used to extract the finger feature,

hand geometry, and palm features. In order to ensure sufficient scanning quality, the user will be required to place their hand on the scanner [4].

Several physiological features can be extracted from the hand and used as biometrics. Of these, the most popular humanoid biometric method is using hand veins, which has characteristic advantages. Since their source lies internal and inaccessible to the body, these are comparatively much more secure and difficult to forge. Furthermore, the capturing of vein patterns is possible only when a subject is alive, thus excluding the possibility of anyone stealing such patterns from a dead body [5]. Moreover, vein patterns remain the same throughout life and do not vary with age. Hand images also use contactless sensors for capturing images, thus consuming less hygiene and proving to be convenient. Research has also shown that vein recognition is likely to be unaffected by different types of distortion like motion blur, defocus, aging of sensors, and compression [6]. Traditional systems do not identify the legitimate person from the intruder; by which one can access the system fraudulently Biometric systems do offer a greater deal of convenience as they can be used without having to memorize passwords for unlike accounts with access granted through an individual's single biometric trait. Biometric systems provide great convenience, especially when compared with traditional systems, but are not immune to attack [7].

Related Work

D. Tejaswi, et. al. (2024) [8], provide an overview of palm vein recognition, highlighting its key characteristics, benefits and applications across various industries including security, healthcare, finance, and personal devices. Additionally, it discusses the role of image processing in enhancing the accuracy and reliability of palm vein recognition systems, as well as future research directions aimed at further improving this technology. Overall, palm vein recognition offers a promising solution for secure and convenient authentication in diverse environments, with the potential to revolutionize the way individuals are identified and verified.

Tuti Sandhya et. al. (2024) [9], explained that palm vein identification relies on unique patterns within the palm veins, illuminated by Near Infrared Light (NIR) with wavelengths from 760 nm to 820 nm, penetrating the skin up to 5mm. The absorption of NIR by deoxygenated blood in veins creates distinct dark patterns. However, this high wavelength light may cause skin and

tissue infection. Vein networks are captured via infrared-sensitive cameras, with captured images preprocessed to remove noise and features extracted for recognition. Feature extraction primarily involves network segmentation, creating reference maps for subsequent recognition. These feature maps serve as blueprints for neural networks, facilitating streamlined identification processes.

Rama Vasantha Adiraju et al.(2021) [10], showed a comprehensive survey on finger and palm vein recognition schemes. The increased use of online data due to the growth of the Internet necessitates enhanced security through unique identification. While passwords and PINs are commonly used for security, the palm vein identification system offers improved security through a unique identification scheme. However, this method tends to be costlier compared to alternative methodologies.

Fawad Ahmad et al. proposed Lightweight and Privacy-Preserving Template Generation for Palm-Vein-Based Human Recognition (2019) [21], aiming to achieve a higher accuracy rate in palm vein recognition schemes. Pre-processing was initially performed on collected datasets, followed by feature extraction. The extracted features were trained to capture the texture, and randomization and quantization methodologies were applied to enhance accuracy. However, the quantization process resulted in a decrease in image quality.

Material and Methods

The proposed system consists of three phases: data processing, Features Extraction, and feature evaluation. Each phase includes multiple processes that work together to fulfill the system's aims. The data set that has been used is HandVein_RL850 [12] which consist of 420 images of hand veins. Figure (2) shows samples of the dataset.

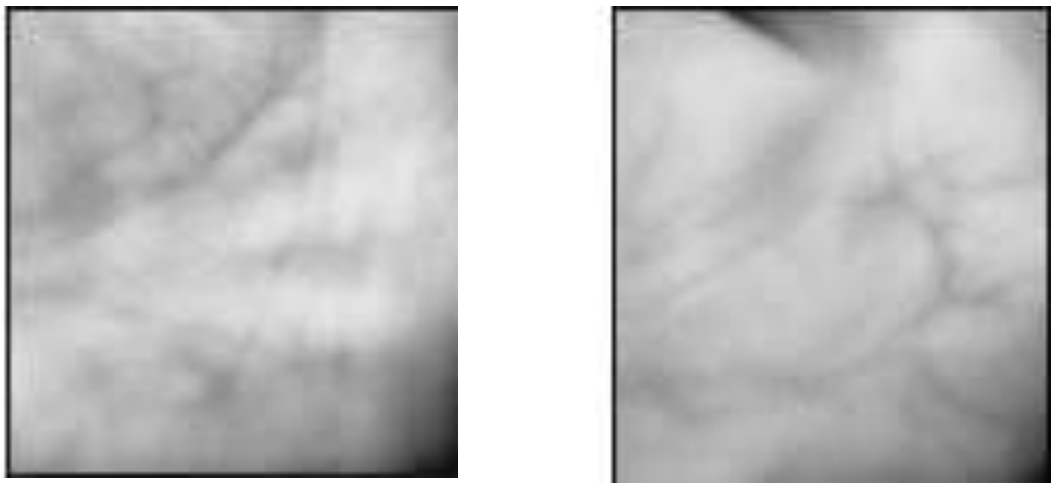

Figure 2: Samples of the Data Set

Hand veins are filtered using box filters, HFEF and CLAHE. Then, the filtered data are discrete cosine transformed (DCT) in order to separate the image into parts with different importance. The DCT transformed data are fused to combine two or more images into one composite image, which integrates the information contained within the individual images. The result is an image that has a higher information content compared to any of the input images. The inverse discrete cosine transform reconstructs a sequence from its discrete cosine transform (DCT) coefficients. The IDCT function is the inverse of the DCT function which is used to merge all the sub-bands image. Data pre-processing can be shown in figure (3 left part). Feature extraction is done depending using PCA NET. Feature extraction is shown in figure (3 right part). However, an algorithm (1) clarifies data pre-processing and feature extraction.

Academic Science Journal

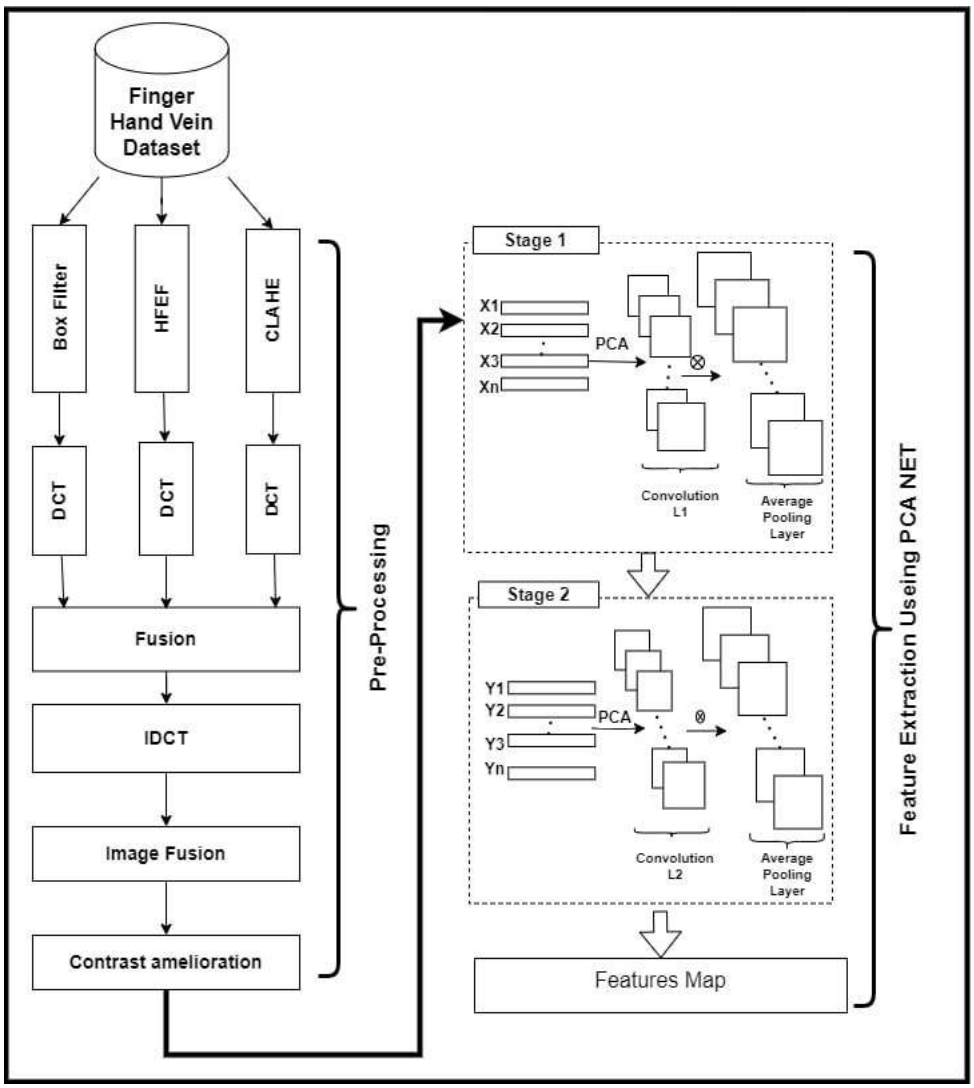

Figure 3 Block diagram of preprocessing and feature extraction

Algorithm (1): Preprocessing and Feature Extraction Algorithm
Input: Hand Vein Image Output: Extracted Features
Begin Step 1: input image data set. Step 2: implement Box filter , HFEF , CLAHE Step 3: implement DCT Step 4: implement fusion Step 5: implement IDCT Step 6: implement image fusion Step 7: contrast amelioration Step 8: feature extraction using PCA NET END

Evaluation of extracted features is very important in this work, because the features in the attack phase need to be compared later, which represents the success of the poisoning process. Therefore, four types of classification are used in this evaluation.

A. SVM Classifier

The procedures for the Support Vector Machine (SVM) are presented in algorithm (2).

Algorithm (2): SVM Algorithm
Input: Preprocessed data Output: Evaluated features
Begin Step 1: Load the preprocessed data. Step 2: Split the dataset for training and testing n-1. Step 3: Calculate classification function. Step 4: Find the hyper plan. Step 5: Find optimal hyper plan. Step 6: Split the trained dataset from the SVC for training and testing n-1. Step 7: Fit a linear regression to predict the targets from covariates for all preprocessed data. Step 8 : Find the accuracy. Step 9: Return Accuracy. END

B. Logistic Regression Classifier

The procedure of logistic regression is: first, to get the minimum value, the maximum value and the step length from the first column of the two-dimensional array to generate a new array. Then, get the maximum value and the step length from the second column of the two- dimensional array to generate another new array. Logistic regression method is shown in Algorithm (3).

Algorithm (3): Logistic Regression
Input: preprocessed data Output: Evaluated features
Step 1: load pre-process data. Step 2: Get binary outcome. Step 3: Get dependent variable vectors. Step 4: Find model's parameters vector. Step 5: Find the accuracy. Step 6: Return Accuracy. END

C. Naive Bayes Classifier

The Naive Bayes algorithm is clarified in algorithm (4).

Algorithm (4): The Naive Bayes Algorithm
Input: Preprocessed data. Output: Evaluated features.
Begin Step 1: Represent the data as a vectors. Step 2: Classify the vectors. Step 3: Calculate the probability. Step 4: Approximate the probability of continuous data. Step 5: Find the accuracy. Step 6: Return Accuracy. END

Results

In image processing, entropy is a metric that quantifies the quantity of information or uncertainty included in an image. Higher entropy indicates more depth or diversity in the image, whereas lower entropy indicates more uniform or simpler images. It measures the complexity or unpredictability of pixel intensity distributions. In order to comprehend and control the quantity of information included in an image, entropy is frequently utilized in image compression, segmentation, and analysis. The output of feature extraction by using PCA net is shown in Figure 4

Academic Science Journal

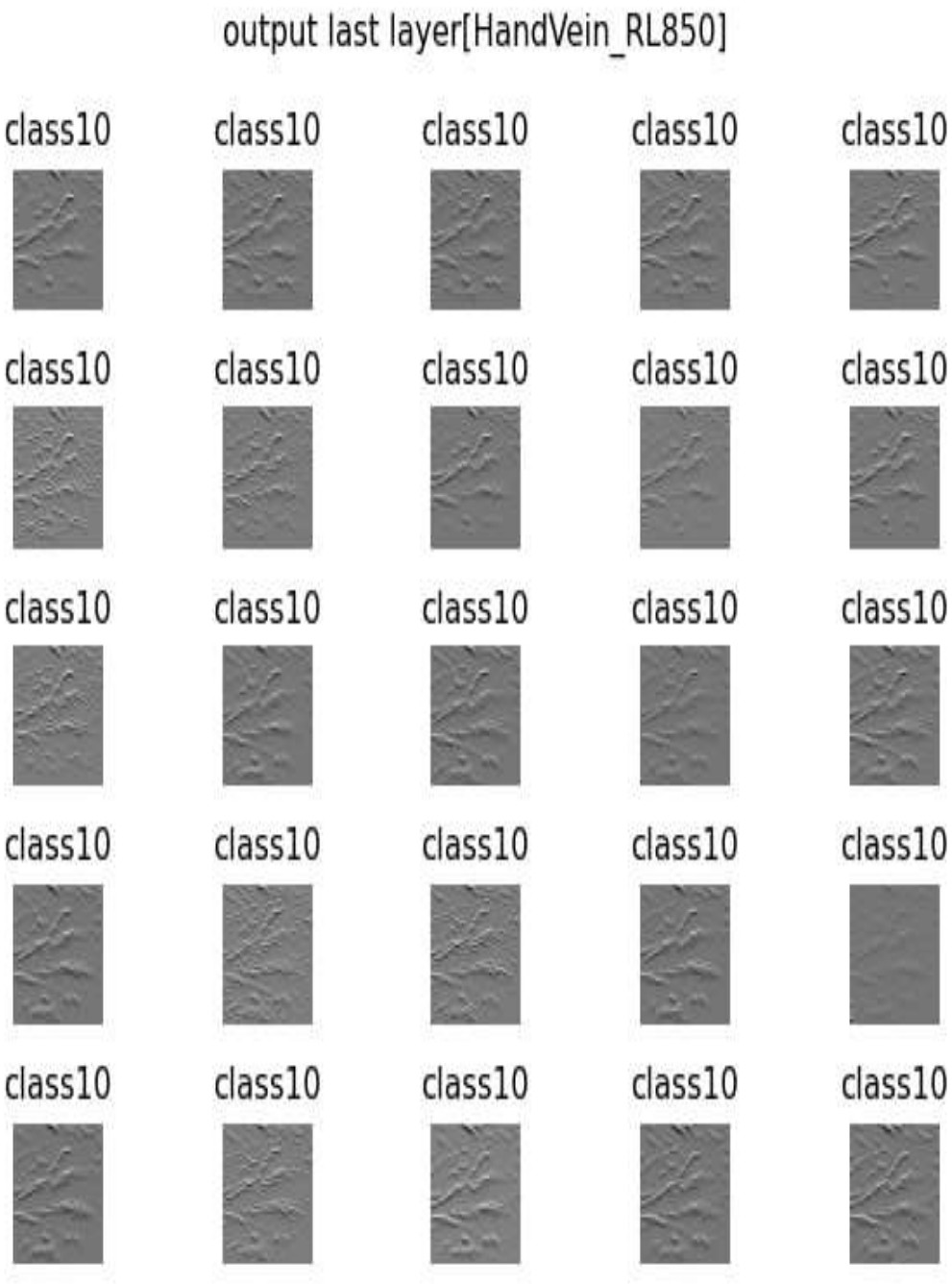

Figure 4: Hand Vein Classes

The number of features that were extractor is 717 as shown in Figure (5).

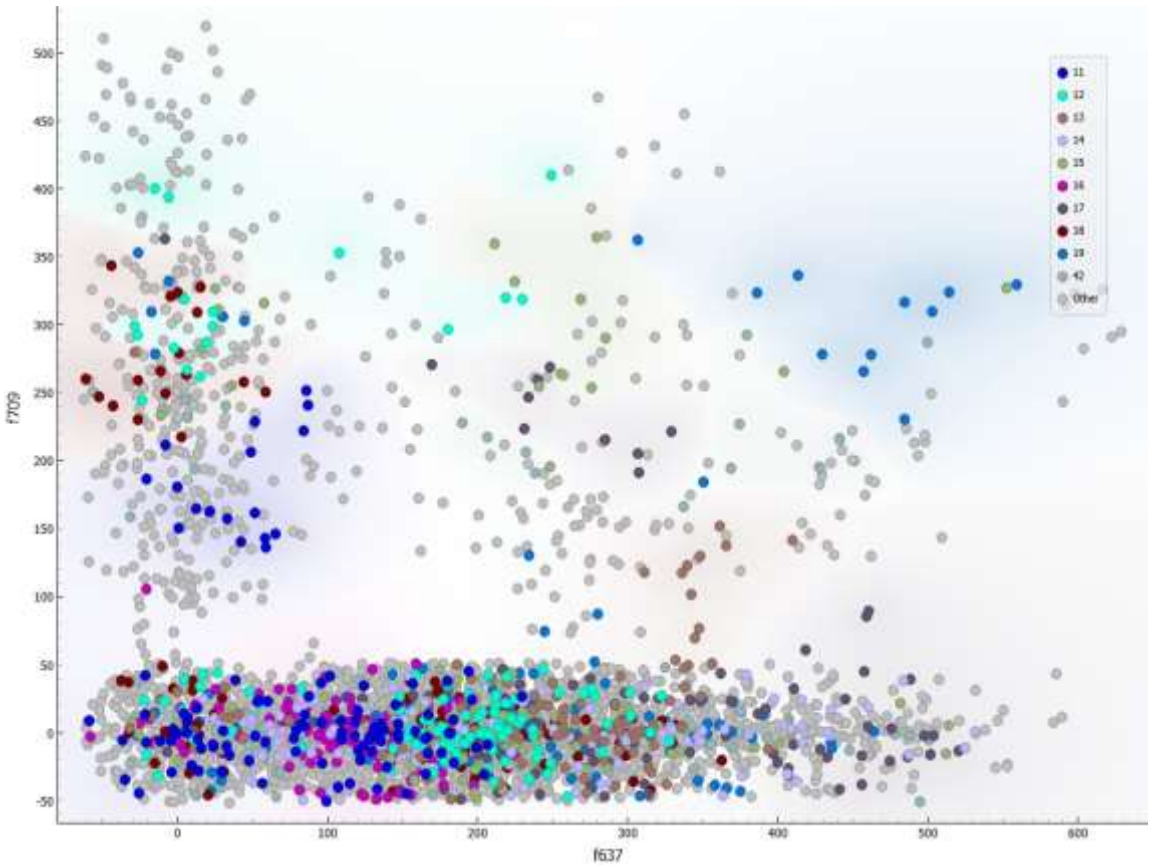

Figure 5: Sample of Features Extracted

The third stage is to test the features and verify the accuracy of the classification, which is the evaluation. Table 1 shows the evaluation measurement for the algorithms.

Table 1: Algorithm Evaluation Metric

Algorithm	ACC	AUC	F1	PRE	Recall	MCC
Logistic regression	99.71	96.0	96.1	96.1	96.0	95.9
SVM	99.6	96.0	96.1	96.2	96.0	95.9
Random Forest	98.0	88.2	88.2	88.3	88.2	87.9
Naïve bayes	91.0	48.6	50.7	48.6	48.6	47.7

ACC (accuracy), AUC (Area Under Curve), F1 (F1-Score), PRE (Precision), Recall, and MCC (Matthews Correlation Coefficient) are among the measures used to summarize the performance of four algorithms (Logistic Regression, SVM, Random Forest, and Naïve Bayes) in the table. This is an explanation and breakdown:

1. Logistic Regression Excellent separability between classes is shown by an ACC of 99.71. AUC: of 96.0%. F1, Recall, and Precision: All at roughly 96%, indicating a well-rounded performance. MCC: 95.9%: Predictions and true labels have a strong association.

   With consistently good metrics across the board, logistic regression works incredibly well with this dataset. Its effective linear decision boundary suggests that linear separation is a good fit for the problem.
2. Support Vector Machine (SVM) ACC: 99.6 – Excellent but marginally behind Logistic Regression. Recall, AUC, and F1 are all 96%, which is consistent with logistic regression. MCC: 95.9%; Logistic Regression is just as strong, with consistently good metrics across the board, logistic regression works incredibly well with this dataset. Its effective linear decision boundary suggests that linear separation is a good fit for the problem.
3. Random Forest's ACC of 98.0 is lower than that of SVM and Logistic Regression. AUC, F1, Precision, and Recall: Approximately 88%, indicating respectable but unimpressive performance. A respectable connection between predictions and true labels is indicated by the MCC of 87.9%. While Random Forest does reasonably well, it is not as good as SVM and Logistic Regression. In this dataset, it may have trouble with overfitting or miss some intricate patterns. Its performance might be enhanced via hyperparameter adjustment, such as changing the maximum depth or the number of trees.
4. Bayes Naïve: The algorithm with the lowest ACC of all is 91.0. Very poor classification accuracy and (AUC: 48.6%). F1, precision, recall: 48–50%, this suggests inadequate performance and balance. MCC: 47.7%: Predictions and true labels have a weak association. On this dataset, Naïve Bayes performs poorly. Among the most plausible causes are: Naïve Bayes makes the assumption that features are independent, but this may not be the case for this dataset. Class Imbalance: Minority class samples may be incorrectly classified by Naïve Bayes if the dataset is unbalanced. Zero-Frequency Problem: Without smoothing methods like Laplace smoothing, Naïve Bayes performs poorly if specific feature-class combinations are missing from the training

data. With nearly similar metrics, SVM and logistic regression are both excellent performers. SVM might do better with non-linear kernels if the data is more complicated, but logistic regression is easier to use and understand. Although Random Forest is less competitive, it does fairly well, with accuracy and other measures hovering around 88%. Tuning the hyperparameters might help. Naïve Bayes performs noticeably worse. This implies that the dataset contains intricate feature dependencies that go against the independence requirement of Naïve Bayes. It might also point to issues like unequal class distribution or the need for improved preprocessing.

Conclusion

This study demonstrated hand vein patterns to be a secure and reliable biometric feature for authentication systems. Advanced image processing techniques combined with feature extraction using PCA Net proved effective in capturing the critical vein attributes. Among the classifiers, Logistic Regression and SVM gave superior performance in the machine learning-based classification of hand vein features by providing quite high accuracy and robustness. The results have brought into focus the need for the selection of appropriate algorithms for biometric security systems to enhance authentication accuracy and make systems resistant to identity fraud. Though Naive Bayes had a poor classification performance, the overall system presents a sound framework for secure biometric applications. Further integration of more biometric traits and optimization of machine learning models could be a future direction of research to enhance the performance and adaptability of the system in varied environments.

Source of funding: The source of funding is self-funding.

Conflict of interest: There is no conflict of interest for authors.

References

[1] B. A. Omonayajo, Advancements in Biometric Technology, Near East University, M.Sc. Thesis, (2021)

[2] S. S. Harakannanavar, P. C. Renukamurthy, and K. B. Raja, Comprehensive study of biometric authentication systems, challenges and future trends, Int. J. Adv. Netw. Appl., 10(4), 3958–3968(2019), DOI(10.35444/IJANA.2019.10048)

[3] L. J. Gonzalez-Soler, D. K. Jónsdóttir, C. Rathgeb, and D. Fischer, Information Fusion and Hand Alignment to improve Hand Recognition in Forensic Scenarios, IEEE Access, 2024, DOI(10.1109/ACCESS.2024.3386955)

[4] S. Barra, M. DeMarsico, M. Nappi, F. Narducci, and D.l Riccio, A hand-based biometric system in visible light for mobile environments, Information Sciences, 479, 472-485(2019), DOI(https://doi.org/10.1016/j.ins.2018.01.010)

[5] J. B. Hill, and N. E. Marion, Introduction to cybercrime: computer crimes, laws, and policing in the 21st century, PSI textbook, (Bloomsbury Publishing USA, 2016)

[6] H. T. Luong, H. D. Phan, D. Van Chu, V. Q. Nguyen, K. T. Le, and L. T. Hoang, Understanding Cybercrimes in Vietnam: From Leading-Point Provisions to Legislative System and Law Enforcement, International Journal Cyber Criminology, 13(2), 2019, DOI(10.5281/zenodo.4766538)

[7] A. O. Alaswad, A. H. Montaser, and F. E. Mohamad, Vulnerabilities of biometric authentication threats and countermeasures, International Journal of Information Computation Technology, 4(10), 947–958(2014)

[8] D. Tejaswi, K. Tejasri, K. Sai Supriya, Ch. Sireesha, Palm Vein Recognition Using Image Processing , International Advanced Research Journal in Science, Engineering and Technology, 11(3), March 2024, DOI(10.17148/IARJSET.2024.11330)

[9] T. Sandhya, Gogula S. Reddy, V. Lakshmi, S. Ahuja, Palm Vein Recognition Using Networking, International Conference on Multidisciplinary Research and Sustainable Development (ICMED 2024), ATEC Web of Conferences 392.

[10] A. Rama Vasantha, An extensive survey on finger and palm vein recognition system, Materials Today: Proceedings, 45, Part2, 1804-1808(2021), DOI(https://doi.org/10.1016/j.matpr.2020.08.742)

[11] F. Ahmad, Lee-Ming Cheng, and A. Khan, Lightweight and privacy-preserving template generation for palm-vein-based human recognition, IEEE Transactions on Information Forensics and Security, 15, 184-194(2019), DOI(10.1109/tifs.2019.2917156) .

[12] PLUSVein-Contactless Finger and Hand Vein Data Set, publicly available for research purposes, (http://www.wavelab.at/sources/PLUSVein-Contactless/), accessed at 9/1/2024.