Received: 19 July 2024 | Revised: 12 December 2024 | Accepted: 28 February 2025 | Published online: 23 April 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2025, Vol. 3(3) 295–304
DOI: 10.47852/bonviewAIA52023875

BON VIEW PUBLISHING

EN-IRAM: Enhance Iris Authentication Model with Soft-Vote Ensemble Learner and Gradual Feature Selector

Folu Ajayi[1,*], Adewale Aromolaran[1], Alomaja Victor[1], Al-Amin Usman[1] and Femi Johnson[2]

[1]*Computer Science Department, Yaba College of Technology, Nigeria*

[2]*Department of Computer Science, Federal University of Agriculture, Nigeria*

Abstract: Authentication systems have facilitated the effective safeguarding of sensitive data and information from unauthorized access. To establish a more robust security architecture for user authentication and verification across various systems, concerted efforts have been focused on integrating one or more biometric characteristics with encryption techniques or machine learning frameworks. Despite the excellent results achieved by these methodologies, a persistent dilemma persists in ascertaining the necessary quantity of features essential to furnish systems with sufficient security safeguards. This paper investigates the prospective benefits of utilizing the gradual feature space reduction algorithm and Vote Ensemble Learner in relation to iris features, along with their impacts on the effectiveness of developing reliable and efficient user verification systems. The assessment outcomes generated by the proposed enhanced iris authentication model across a diverse array of datasets demonstrate accuracy, precision, and *F*1-score ranges of 0.77 to 1.00. The developed model achieves an average score of 0.01%, 0%, and 0.03% for EER, FAR, and FRR respectively with an accuracy of 99.99%.

Keywords: authentication, vote ensemble, feature selection, privacy, security access, machine learning

1. Introduction

The security of data and information has been an evolving process in the computing field. It has evolved through the years into diversifying computing security into related sub-fields such as cyber security, information management, and IT infrastructure safety [1]. Information and computer security processes are required to protect valuable data and information both online and offline. Many authentication systems are being developed as proactive measures to forestall and prevent unauthorized access to services or data [2]. Besides what users know and provide during authentication phases such as username and password [3] with other security measures, the encapsulation of the users' biological traits is also being utilized in more advanced systems [4]. Common biometric features used for authentication systems include fingerprint scans, facial or retina scans, and voice recognition authentication systems [5, 6]. Individual biometrics is rated for their uniqueness in identifying unique users and is proven to be more efficient than using tokens or hardware items.

The performance metrics of biometrics systems have also been proven to surpass multi-factor cryptographic authentication systems for access control [7, 8]. Selected distinct human attributes utilized have helped to discourage and prevent impersonation, identity theft, or mismatch. Authenticated users are granted access to system's resources (data or services) upon a successful match of extracted biological attributes with those in the database.

*Corresponding author: Folu Ajayi, Computer Science Department, Yaba College of Technology, Nigeria. Email: abiodun.ajayi@yabatech.edu.ng

Adopting the iris, a part of the human eye for user authentication is promising for lofty performance measure in authentication and verification systems.

The iris feature of the human eyes consists of over two hundred and fifty (250) unique elements, which describe human identity through numeric feature vectors [9]. Studies have shown that the feature vectors are unique (even in the case of twins), they can be introduced to a classification algorithm for modeling [10, 11] before deployment in user authentication systems, and the resulting feature vector data are also suitable for a data mining technique combined with ensemble learning for developing users authentication systems. In addition towards improving the performance of a verification and authentication system, this adoption ensures a reduction in verification failure [12].

Furthermore, identified research gaps have necessitated the need to enhance the authentication processes. Noticeable gaps reflect the absence of optimal feature selection task for implementation [13], the vulnerability of the biometric system with no feature selection mechanism [14], non-deployment of feature selection in data processing phase and methodology [15, 16], as well as the existence of the vulnerability of the single factor system [17].

In this paper, an enhanced iris authentication model is developed to bridge the highlighted gaps by deploying a modified soft-vote ensemble machine learning technique with a gradual feature selection reduction approach (GFSRA) on well-trained biological comprising of extracted iris feature vectors from proposed system authenticated users.

The contributions of this study lie in its methodology which become imperative for efficient data handling especially with large data samples for pattern recognition which are listed below.

1) The incorporation of GFSRA which contributes to enhancing the performance of the system.
2) The memory of the system is efficiently utilized by eliminating redundant features and selecting the most optimal subset dataset.
3) The computational cost of the system is reduced as system utilizes only numeric transforms of extracted iris images.
4) The utilization of the vote ensemble model, comprising support vector machine (SVM), multilayer perceptron (MLP), and decision tree (DT) base learners, demonstrates the efficacy of combining diverse learning strategies for improved user authentication.

2. Literature Review

Academics are advancing and refining authentication systems at an expedited pace in response to the increasing demand for secure and reliable user verification methodologies. Researchers have examined a variety of strategies and methodologies to enhance authentication procedures [18, 19]. These encompass both conventional techniques and sophisticated machine learning methodologies. They include a spectrum of authentication mechanisms, such as biometric authentication systems [20], location-based verification, one-time password frameworks, multi-factor authentication, and conventional password-based authentication [16, 21, 22].

As a growing number of biometric systems depend on passwords for verification, the financial and computational burdens associated with storage solutions become notably onerous, particularly as user demographics expand. The investigation conducted by Sandhya et al. [23] presents a solution through the introduction of a multi-instance cancelable iris system that preserves the feature vector as a cancelable template. This pioneering system utilizes a triplet loss function for feature extraction through a comprehensively trained convolutional neural network (CNN), thereby facilitating pattern comparison, which in turn reduces access latency while simultaneously improving both security and precision.

Although it has been demonstrated that using an individual's unique physiological or behavioral characteristics as a verification or authentication process is more reliable [24, 25] than depending on a shared secret or key, biometric system authentication is not without its vulnerabilities, such as accuracy-related problems (false positive and negative matches, e.g.,) and potential biometric system attacks [26]. An earlier attempt at authentication using mobile devices is recorded [15]. They implement a multi-modal authentication system with smartphones. Utilizing face, periocular, and iris biometric data points, a multi-modal biometric identification system was developed in their study. Two Samsung mobile devices, the Galaxy S5 and Galaxy Note 10.1, are used to test the suggested methods. Scale Invariant Feature Transform, Binarized Statistical Image Features, and Speeded-Up Robust Features are among the feature extraction methods used in the investigation.

A secret-sketch graphical authentication framework (SEC-SKETCH) is conceived by Joseph et al. [13]. The framework of the SEC-SKETCH synergistically integrates a username, a textual passcode, and a recall-based image sketch methodology to augment security measures. Analytical modalities such as threshold and percentage accuracy, which confer significant resistance to both hidden-camera and shoulder-surfing attacks, are incorporated into the evaluative framework. Upon comparative performance assessment with analogous methodologies, it was determined that the SEC-SKETCH outperformed its counterparts, attaining scores of 0.15%, 0.02%, and 0.10% for false match rate, equal error rate (EER), and false non-match rate, respectively.

Vazquez-Fernandez and Gonzalez-Jimenez [27] also test a mobile device-based biometric system on a relative training image set from repositories for usability, robustness, security, and mitigation against spoofing attacks. The work of Ackerson et al. [28] explores the use of recurrent neural network to identify irregular patterns from the performance analysis on sequential data useful for recognizing human expressions. They also propose various applications of their model in detecting flight conditions and assessing the lifespan of turbine engines. FIDO and Public Key Infrastructure (PKI) with no feature selection mechanism are used in the design of a secure biometric authentication system for testing in finance environments [29]. The authors compared already registered biometrics with new ones, and PKI was used to address the issue of user authentication on several web platforms in financial technology (fintech) business use cases. Their study further revealed that a centralized biometric framework for user authentication will aid in safeguarding users' privacy and improve consumer convenience. The efficiency of deep learning techniques for better results is investigated by Ibtehaz et al. [14] through the development of EDITH, a deep learning framework that uses four open data sources containing ECG signals. These signals are inputted into the developed network. The Siamese architectural design of the model is such that it reduces the EER for improved accuracy. Their study also incorporates a dual method for identity verification that depends on retrained data of new intakes and closed environment identification of a fixed number of participants. Despite the robust enrollment and evaluation dataset, the system is noted to be vulnerable due to the redundancy of the dataset. In similar research conducted by some researchers, system vulnerability is detected due to lack of an authentication driving module in the model proposed by researchers [30].

In an attempt to optimize the quality of data and reduce the complexity of authentication systems while still generating an accurate result with improved security features, Ying and Nayak [31] propose a lightweight, untraceable remote user authentication system for multi-server 5G networks. They apply elliptic curves and self-certified public key cryptography to confirm the legitimacy of users and servers on a tamper-proof system with the discrete logarithm encryption technique. Other researchers have deployed machine learning feature extraction techniques [7] for biometric data extraction towards developing highly secured authentication systems [32].

The implementation of iris recognition systems is demonstrated in the research conducted by Ali and Shaker [20]. They introduce a secure electronic voting mechanism that enables users to be authenticated through their iris-derived characteristics. The data acquired from the images captured are systematically stored, encoded, and subsequently decoded utilizing specialized modified encryption and decryption algorithms prior to transmission. The computational duration for the image-capturing process is also addressed.

A hybrid methodology integrating segmentation and edge detection has been proposed by El-Sofany et al. [33] to bolster data security within cloud systems, following a comprehensive evaluation of the advantages and disadvantages of contemporary techniques applied to three distinct datasets, namely MMU, IITD, and CASIA Iris Interval V4. Aspects including accuracy, robustness, computational complexity, and model optimization, along with their comparative performance, were meticulously scrutinized during the development of their proposed model. They employ CNNs and Hamming distance to assess and extract distinctive features from the aggregated image datasets. The

proposed model achieves recognition accuracies of 0.95, 0.97, and 0.99 on the CASIA, IITD, and MMU iris datasets, respectively, showcasing the different performance of the datasets.

Adamović et al. [34] employ machine learning methodologies and stylometric characteristics implemented on the Weka platform, utilizing R and Python packages, to develop an iris-based authentication framework. The proposed methodology entirely eradicates the false acceptance rates (FAR) typically observed in traditional iris-based systems, without incurring additional false rejection rates (FRR). They formulated two categories of features predicated on linguistic dependency; through the application of the Base64 algorithm, they successfully converted normalized iris segments into a novel form of biometric textual data while preserving its statistical attributes. A straightforward active learning approach designed to mitigate the class imbalance issue is executed by implementing the Synthetic Minority Over-sampling Technique (SMOTE) and the Majority Weighted Minority Over-sampling Technique. The assessment of the model, conducted on the comprehensive CASIA dataset utilizing the Random Forest algorithm, yielded an accuracy rate of 0.9999. Nonetheless, the optimal accuracy of 1.000 is attained through the utilization of the SMOTE technique.

Despite the multitude of studies from existing literature with a wide array of feature selection techniques [28, 35], the implications and advantages of the GFSRA spaces, wherein subsets of features are systematically selected, remain under-explored. This paper aims to bridge this gap by investigating the effectiveness of GFSRA in enhancing user authentication models. The remaining portions of this paper are organized as follows: In Section 3, the process for developing the improved biometric (iris) authentication model is described, including the implementation methods, and outcomes are laid out in Section 4. Finally, the Sections 4 and 5 give the discussion and conclusion, respectively.

3. Research Framework and Methodology

The model's framework authentication mechanism is a four-phase biometrics-based user verification system that uses extracted numeric features from iris pattern recognition attributes to train a machine learning model as depicted in Figure 1. The training set consists of iris images of authentic users for easy recognition, which is then fitted by a Vote Ensemble Learner algorithm to build the artificial intelligent authentication system. Feature vectors are extracted from the input data through knowledge transfer technology of image embedding functionality and are consequently deployed for the ensuing phase of feature engineering. For the subsequent machine learning stage, the feature engineering result constitutes the optimal training set. The suggested intelligent model, which is the result of supervised machine learning training, has the ability to identify biological patterns from a user who plans to obtain access through the machine learning testing phase. Each user is identified through their names, thereby alienating a strange user. At the fourth stage of the conceptual framework, access is either permitted or refused based on the authentication result. The methodology is described as follows.

3.1. Data acquisition and image processing

The data utilized for this research are collected from Kaggle, a public data repository. The downloaded dataset comprises two image folders containing well-classified captured images of 103 authorized users. Every authorized user has twelve (12) unique photos that depict various facial features. The obtained dataset has 1236 picture occurrences in total. The image signals are preprocessed to extract only the iris of each instance represented in the dataset. The image cropping was achieved on the Windows operating system. Thereafter, the needed image signals are collected and saved in a folder, comprising the iris features of each instance.

3.2. SqueezeNet encoding and feature extraction

The input iris image signals are subjected to image encoding and automated extraction of a thousand (1000) genetic numeric features from each of the input image with the SqueezeNet neural network. We also apply normalization technique to foster coherence within the generated features presented in a CSV file, for the graduated feature selection exercise to be performed. Figure 2 depicts the squeezeNet encoding and extraction phase as performed in the Orange data mining toolkit.

Figure 1
The framework of the developed enhanced Iris-based authentication model

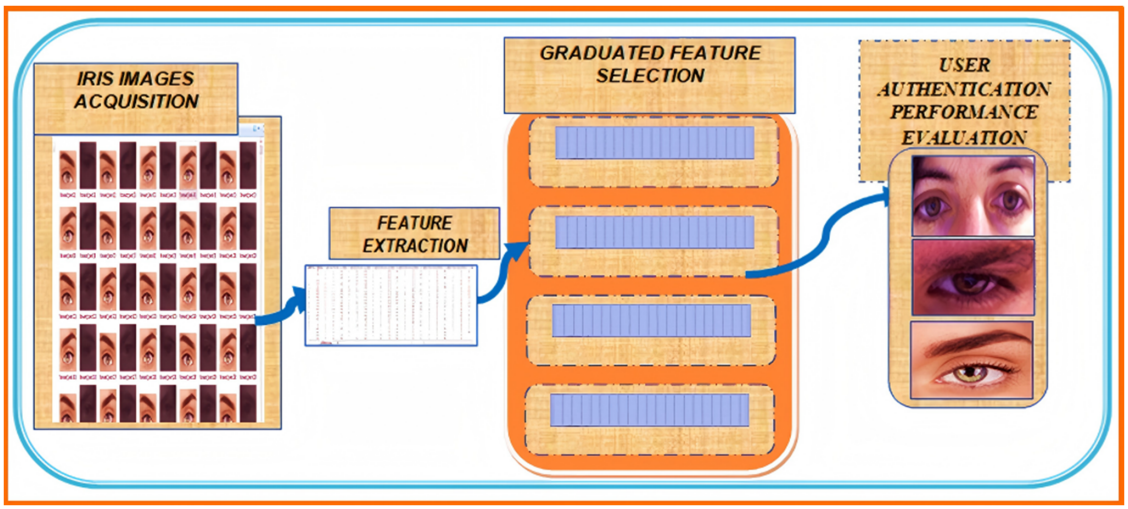


Figure 2
Iris image embedding and feature extraction

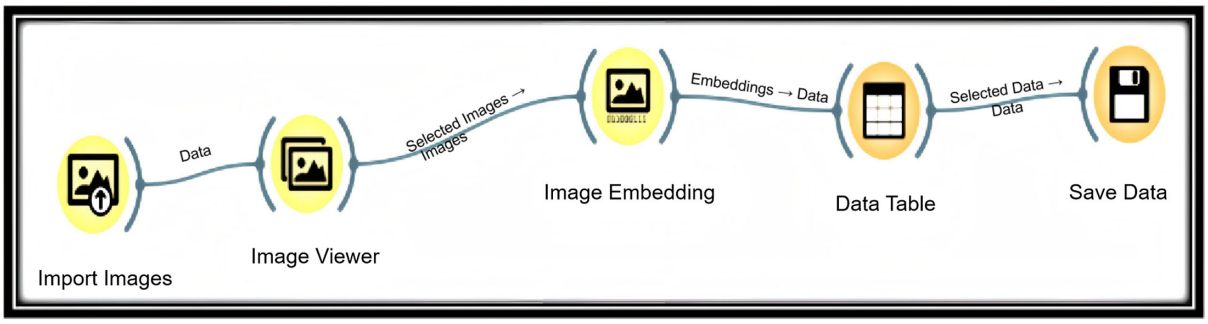


Table 1
Distribution of graduated features

S/N	Attribute range	Class category	Number of features
1.	A_i	Class 0	1000
2.	A_1	Class 1	800
3.	A_2	Class 2	600
4.	A_3	Class 3	400
5.	A_4	Class 4	100
6.	A_5	Class 5	10

3.3. Feature selection using graduated selector

The graduated feature selection approach is deployed to reduce the features needed for training the model and also helps the model adapt to the decreasing feature dimensions by balancing the model's complexity with its predictive performance.

A group of five different files named (A1–A5) are initially created and marked as subsets to hold reduced numeric iris-based features in the proportions 800, 600, 400, 100, and 10 respectively as described in Table 1. Each subset is computationally valued using principal component analysis to represent the attributes.

The extracted numeric features of the iris dataset are also graduated in the following subsets of 800 attributes (a1), 600 attributes (a2), 400 attributes (a3), 100 attributes (a4), and 10 attributes (a5). The mathematical equations for the sequence of feature selection are represented in Equations (1)–(7), respectively. Equation (1) represents the general SqueezeNet formula for encoding and extraction of the iris-based dataset. Equation (2) describes the entire features from which smaller subsets (a_1, a_2, a_3, a_4, a_5) can be derived. The derived featured subsets are represented in Equations (3)–(7).

$$A_i = \sum_{i=1}^{m} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ n = 1 \ldots 1000 \tag{1}$$

$$a_j \subset \sum_{i=1}^{1000} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ j = 1 \ldots n$$

$$\text{such that} \quad \exists \ a_1, a_2, \ a_3, a_4, a_5 \in A_i \ . \ . \ . \tag{2}$$

Where:

$$a_1 = \sum_{i=1}^{n} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ n = 1 \ldots 800 \tag{3}$$

$$a_2 = \sum_{i=1}^{p} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ p = 1 \ldots 600 \tag{4}$$

$$a_3 = \sum_{i=1}^{q} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ q = 1 \ldots 400 \tag{5}$$

$$a_4 = \sum_{i=1}^{r} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ r = 1 \ldots 100 \tag{6}$$

$$a_5 = \sum_{i=1}^{t} [\text{SqueezeNet}_{\text{Features}}]_i \ \forall \ t = 1 \ldots 10 \tag{7}$$

3.4. Model training and evaluation

For the purpose of experimental training, an ensemble of Soft-vote techniques with Sequential Minimal Optimization (SMO) Algorithm, multilayer perception (MLP), and DT as base classifiers are used as shown in Figure 3. This is to provide an improved accuracy, reduced overfitting, improved generalization and efficiency of the model. Each learner is also trained, validated, and tested with varying proportion of samples deployed through a collaborative approach for a robust final binary classification of authorized users. The soft-vote technique deployed in this study then combines and computes the average probabilities of the base classifiers concerning their individual classification decision and the resulting decision is returned as the majority vote.

Algorithm 1 delineates the procedural framework for the optimization process employed by the SMO during the training phase of the model. The Python programming language is used to carry out the implementation. Modified ensemble training is performed at 10-fold cross-validation on the collected dataset utilizing soft-vote function in python programming language. The function after training the model also performs validation generating evaluating performance metrics such as accuracy, precision, *F*1 score, and ROC or AUC on a new list of independent variable ranges.

3.5. GUI user authentication

This phase implements a graphical user interface (GUI) for the vote ensemble model. All the necessary libraries, including tkinter

Figure 3
Developed model's training interface

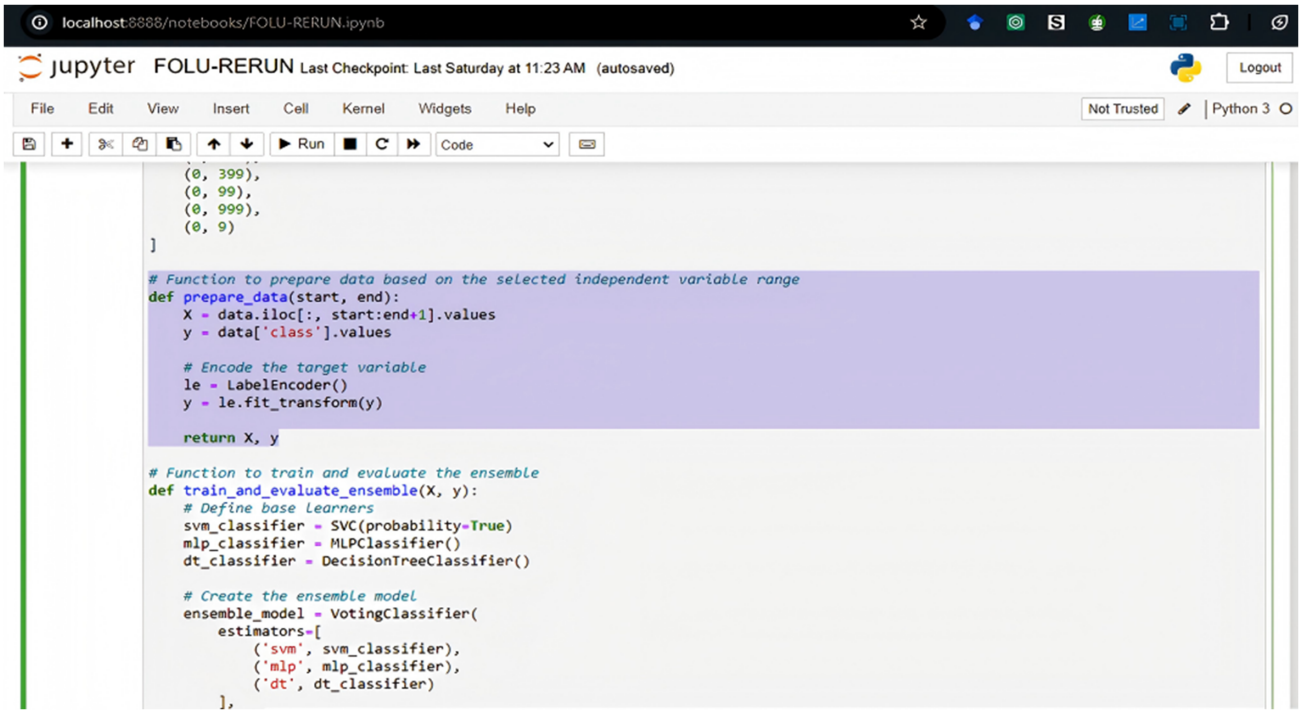
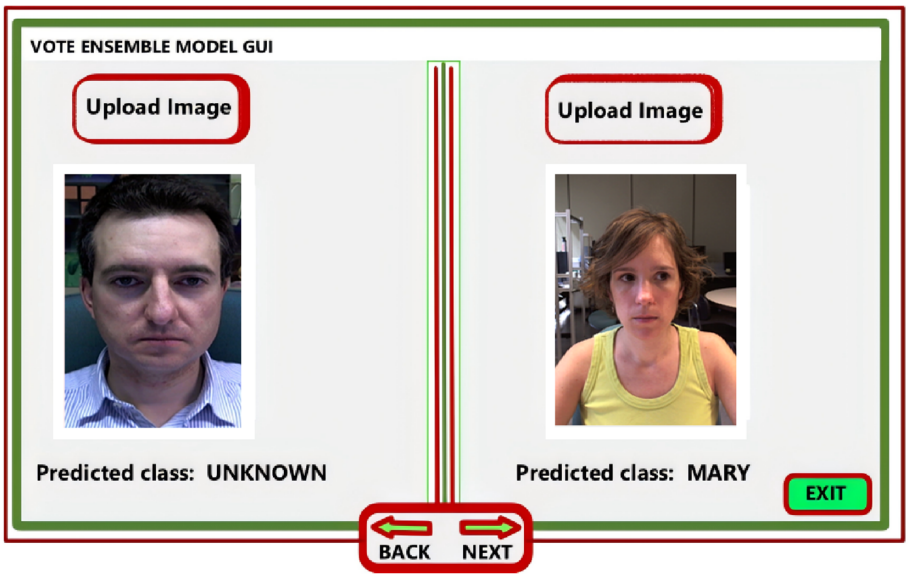


Figure 4
Vote ensemble GUI interface for user authentication



(Tk) for GUI, PIL for image processing, Tensor Flow for machine learning, and various components from Sklearn with metrics for performance evaluation are imported. The main GUI window is created using Tk, and various GUI elements such as buttons, labels, and image display areas are configured appropriately as depicted in Figure 4. The upload-image function is assigned to the "Upload Image" button, and labels are created to display the predicted class, accuracy, precision, recall, and *F*1 score. The update metrics function calculates performance (accuracy, precision, recall, *F*1 score) and updates the corresponding labels on the GUI.

Algorithm 1: The Sequential Minimal Optimization

Input: Ground truth labels (p_t, q_t), t = 1, . . . , n, and a small constant f.
Output: optimal output ⊖.
i ← − 1; j ← − 1
∇˜ f(⊖) ← 0
while ⊖ is not optimal **do**
Unsystematically permute samples.
for t ← 1 . . . n do
∇˜ f(⊖)t = $w^T p_t$ − q_t
if ∇˜ f(⊖)t<∇˜ f(⊖) − f and t ∈I_{low} **then** ◿I_{low} is defined in (4).
∇˜ f(⊖)i ← ∇˜ f(⊖)t
i ← t
else if ∇˜ f(⊖)t>∇˜ f(⊖) + f and t ∈I_{up} then
◿I_{up} is defined in (4).
∇˜ f(⊖)j ← ∇˜ f(⊖)t
j ← t
end if
if i ≠ − 1 and j ≠ − 1 then
update θ_i and θ_j according to (13)
update w according to (14)
update ∇˜ f(⊖) according to (18)
i ← −1; j← −1;
end if
end for
end while

4. Results and Discussion

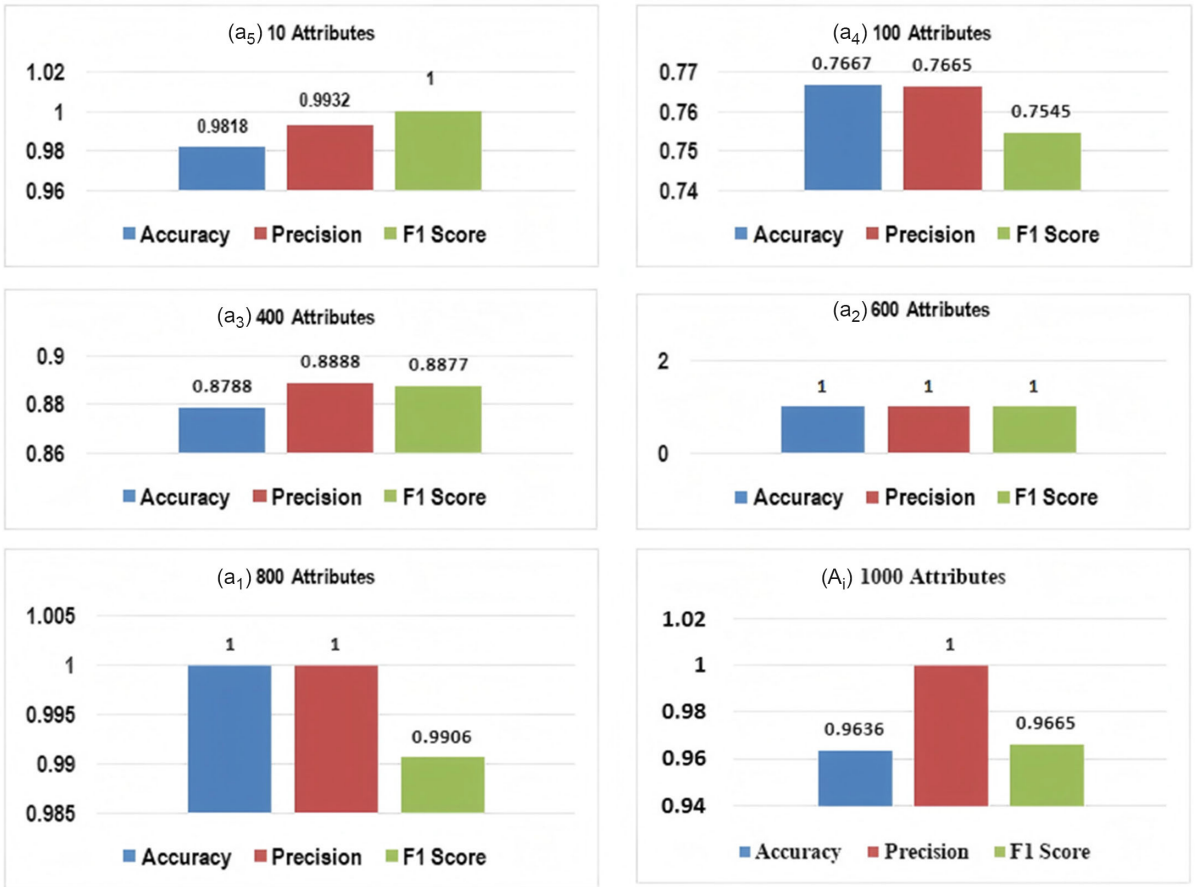
The GUI application is developed utilizing Python programming as illustrated in Figure 4. The development and evaluation processes are conducted on a Lenovo E-450 device equipped with an Intel processor operating at 1.65 GHz, complemented by 4GB of RAM and a 148 GB hard disk, all functioning under the Microsoft Windows 10 operating system. The Anaconda Navigator server is employed alongside the Orange data mining toolbox.

The vote ensemble model is trained on the entirety of the collected images, and the resultant model is preserved for subsequent testing. The Upload Image button facilitates the submission of a user's iris, which will be processed by the initially trained vote model for the purpose of authentication. The vote ensemble is designed to classify the user as one of the legitimate users (designated by their respective names) or categorize them as 'UNKNOWN'.

In addition, the performance of the vote ensemble using the soft-vote ensemble method on the 10-feature attributes subset showed remarkable effectiveness, as evidenced by the high Accuracy, Precision, and *F*1 Score. Each of these metrics attains near-perfect scores with accuracy scaled at 0.9818, precision at 0.9932, and *F*1 Score at 1. This reflects the robustness and harmony achieved by the ensemble model when presented with a highly reduced feature set.

The success of the vote ensemble in achieving outstanding performance with only ten (10) feature attributes can be attributed

Figure 5
Performance chart on attributes selection



to several factors which include the fact that DT algorithm usually excels at capturing complex decision boundaries even with limited attribute numbers, while the SMO is indeed reputable for handling non-linear relationships. The result generated by the improved authentication model at each level where the GFSA is performed is also depicted as shown in Figure 5. An accuracy of 0.9636 at the 1000 attributes selection phase demonstrates the model's capacity to make precise and reliable predictions across a diverse set of attributes.

Furthermore, the examination of precision metrics across different feature vector graduations sheds light on the model's precision at various levels of feature inclusion. Notably, at the 600-feature graduation, the model achieves a perfect precision score of 1, signifying the model's ability to precisely classify positive instances among the selected 600 features. This outcome implies a high level of confidence in the model's positive predictions at this feature subset. As the number of features decreases from 600 to 10, there is a noteworthy increase in precision, reaching 0.9932 at the 10-feature graduation. This emphasizes the model's capability to maintain precision even with a significantly reduced set of features, demonstrating its resilience to feature dimensionality reduction. The progressive decline in precision from 800 to 400 features, reaching 0.8888 at the 400-feature graduation, suggests a trade-off between precision and feature richness.

Considering the *F*1-scores across different feature graduations reveals intriguing insights into the model's performance. At the 600-feature graduation, the *F*1 Score reaches a perfect score of 1, emphasizing the model's exceptional balance between precision and recall when utilizing this feature subset. This suggests that the model successfully identifies positive instances while minimizing false positives, achieving optimal performance at this feature dimensionality. As the number of features decreases to 10, the *F*1 Score remains impressively high at 0.9665, indicating the model's ability to maintain a robust balance between precision and recall even with a significantly reduced set of features. This underscores the model's adaptability to feature dimensionality reduction without compromising its overall performance. However, the *F*1 Score exhibits a slight decline as the feature vector decreases from 800 to 400, reaching 0.9906 at the 800-feature graduation and 0.8877 at the 400-feature graduation. This trend suggests a potential trade-off between achieving high precision and recall, emphasizing the importance of carefully selecting the appropriate feature subset based on the specific requirements of the user authentication system.

Figure 6 depicts the percentage rating of each class category of attributes. AUC values of 0.50, 0.57, 0.42, 0.53, and 0.46 were recorded for the class 1, 2, 3, 4, and 5, respectively. Class 3 and class 2 have the lowest and highest values recorded respectively. The least value indicates a performance close to random chance in distinguishing Class 3 from other classes. Conclusively, among the different feature subsets evaluated, the 100-a4, 600-a2, and 800-a1 attribute sets stand out as the top performers as indicated in Figure 7, showcasing exceptional precision, accuracy, and *F*1 scores. An accuracy of 0.9636 at the 1000 attributes selection phase demonstrates the model's capacity to make precise and reliable predictions across a diverse set of attributes. Figure 8 shows the comparison of the soft-vote ensemble method with similar learners. The results from the chart reflect that the proposed method (soft-vote ensemble) outperforms others in accuracy and precision. This further implies that the model can be deployed for use in iris-based authentication systems.

The accuracy of the proposed system undergoes further assessment utilizing standard biometric authentication value metrics, which encompass the EER, FAR, and FRR. The developed soft-vote ensemble iris-based authentication system

Figure 6
Receiver operating characteristics (ROC-AUC)

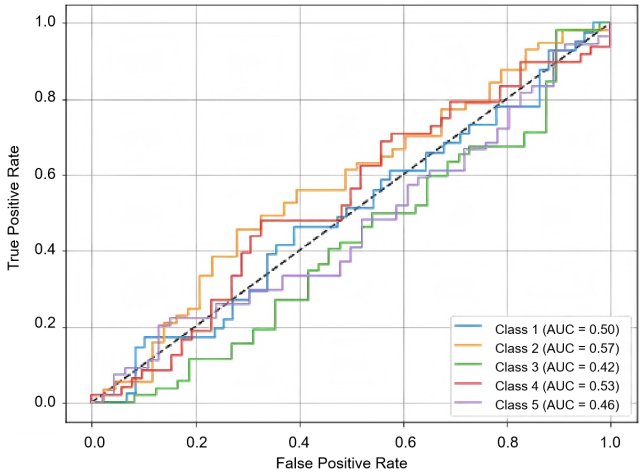


Figure 7
Vote ensemble learner's performance across graduated feature subsets

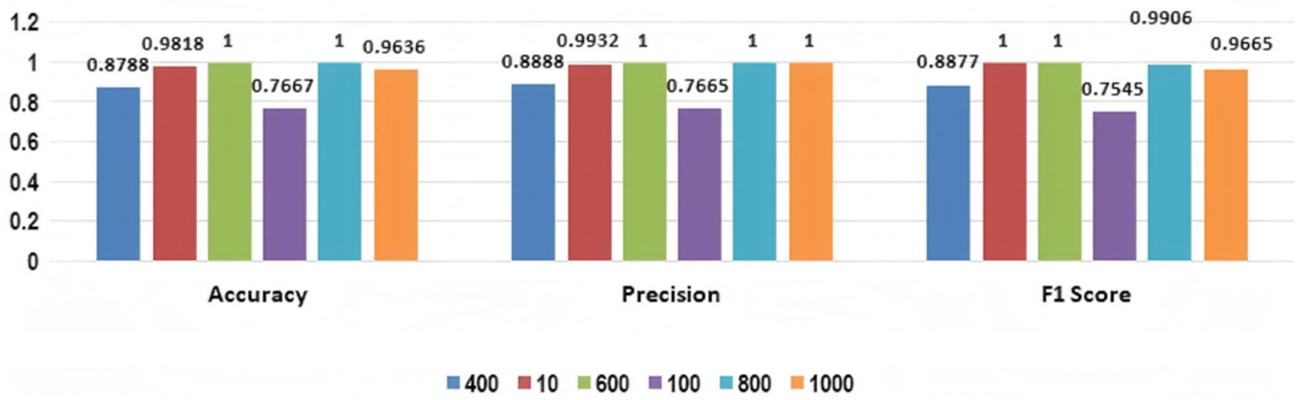


Figure 8
Comparison chart of soft-vote ensemble with other learners

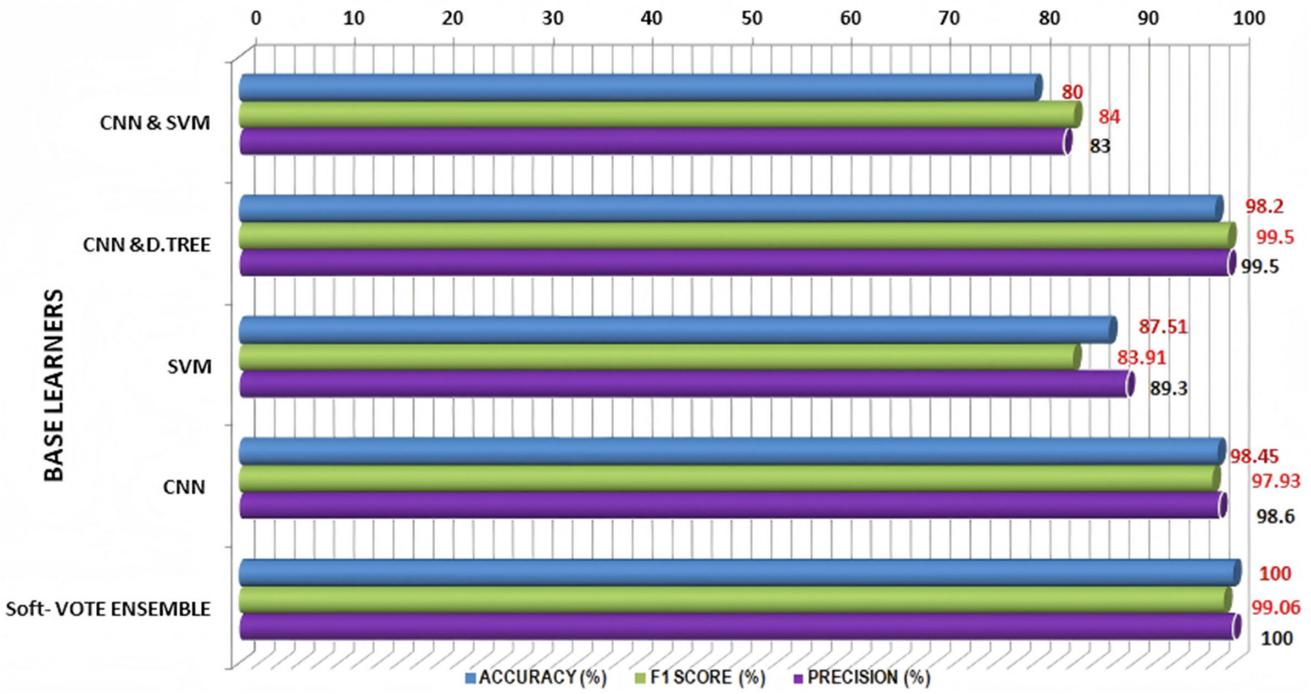


Table 2
Comparison chart of soft-vote ensemble with similar methodology

S/n	Methodology	Accuracy (%)	FAR (%)	FRR (%)	EER (%)
1.	*Hafeez et al.* [36]	99.73	0.12	0.14	0.13
2.	*Joseph et al.* [13]	–	0.15	0.10	0.02
3.	*Tahir and Anghelus [35]*	98.85	0.58	0.58	0.58
4.	*Ullah et al.* [37]	98.2	0.05	0.05	–
5.	*Moi et al. [38]*	98.46%	0.00	1.54	–
6.	*Proposed model (EN-IRAM)*	99.99	0.00	0.03	0.01

achieves an average EER of 0.01%, FAR of 0%, and FRR of 0.03%. Our proposed system also exhibits an accuracy rate of 99.99%, with execution duration of 4.35 s. It is evident from Table 2 that the accuracy of our proposed system surpasses that of alternative methodologies, as the FRR score is minimal and deemed acceptable for applications where security is of paramount importance. Moreover, the system effectively prohibits any unauthorized access, as indicated by the FAR value of zero (0) %, thereby ensuring comprehensive security.

5. Conclusion

This paper accomplished its objectives, which included the deployment of SqueezeNet embedding for feature extraction, the employment of graduated feature selection techniques, and the development of a vote ensemble model for user authentication. The base learners of DT, SMO, and the MLP employed the Soft-vote methodology to approximate the class of the modeling using the 1000 feature attributes returned by the SqueezeNet feature extractor. The feature vectors are clustered into the 1000-Ai subset, 800-a1 subset, 600-a2 subset, 400-a3 subset, 100-a4 subset, and the 10-a5 subset to unravel the user authentication predictive ability of the vote ensemble.

The findings reveal compelling insights into the performance of the vote ensemble model across various feature subsets. Notably, the 100-a4, 600-a2, and 800-a1 attribute subsets emerged as the most reliable and robust configurations, showcasing consistently high precision, accuracy, and *F*1 scores. This outcome underscores the versatility and adaptability of the ensemble model, as it demonstrated exceptional user authentication capabilities across different feature spaces. The 600-a2 subset, in particular, stood out as a pinnacle of performance, achieving perfect scores across all evaluated metrics, indicating an optimal balance between precision and recall.

This observation suggests that the ensemble method, employing DTs, SMO, and MLP models, excelled in handling the complexities of a mid-sized feature space, making it an ideal choice for practical deployment in user authentication scenarios. The 800-a1 and 100-a4 subsets further reaffirmed the ensemble's ability to maintain strong predictive performance across varying feature dimensions, enhancing the model's applicability in real-world biometric systems.

In conclusion, the performance metrics of the proposed model exhibit robustness for its designated application as a biometric

authentication system. The FRR may slightly challenge users' convenience; the absolute security assurance and the overall accuracy remain commendable.

Acknowledgement

The authors acknowledge the efforts of the reviewers of this paper. We also appreciate their meaningful contribution, valuable suggestions, and comments to this paper which helped us in improving the quality of the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in Kaggle dataset, Iris of eye dataset at https://www.kaggle.com/datasets/mohmedmokhtar/iris-of-eye-dataset.

Author Contribution Statement

Folu Ajayi: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Adewale Aromolaran:** Methodology, Formal analysis, Investigation, Resources, Data curation, Supervision, Project administration. **Alomaja Victor:** Validation, Formal analysis, Resources, Data curation, Writing – review & editing, Supervision, Project administration. **Al-Amin Usman:** Software, Formal analysis, Resources, Visualization, Supervision, Project administration. **Femi Johnson:** Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision.

References

[1] Sharma, U., Tomar, P., Ali, S. S., Saxena, N., & Bhadoria, R. S. (2021). Optimized authentication system with high security and privacy. *Electronics*, *10*(4), 458. https://doi.org/10.3390/electronics10040458

[2] Mihajlov, M., Jerman-Blažič, B., & Ilievski, M. (2011). ImagePass-Designing graphical authentication for security. In *7th International Conference on Next Generation Web Services Practices*, 262–267. https://doi.org/10.1109/NWeSP.2011.6088188

[3] Abo-Zahhad, M., Ahmed, S. M., & Abbas, S. N. (2016). A new multi-level approach to EEG based human authentication using eye blinking. *Pattern Recognition Letters*, *82*, 216–225. https://doi.org/10.1016/j.patrec.2015.07.034

[4] Krishnamoorthy, S., Rueda, L., Saad, S., & Elmiligi, H. (2018). Identification of user behavioral biometrics for authentication using keystroke dynamics and machine learning. In *Proceedings of the 2018 2nd International Conference on Biometric Engineering and Applications*, 50–57. https://doi.org/10.1145/3230820.3230829

[5] Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*, *51*, 854–864. https://doi.org/10.1007/s10489-020-01829-7

[6] Rasheed, H. H., Shamini, S. S., Mahmoud, M. A., & Alomari, M. A. (2023). Review of iris segmentation and recognition using deep learning to improve biometric application. *Journal of Intelligent Systems*, *32*(1), 20230139. https://doi.org/10.1515/jisys-2023-0139

[7] Balashanmugam, T., Sengottaiyan, K., Kulandairaj, M. S., & Dang, H. (2023). An effective model for the iris regional characteristics and classification using deep learning Alex network. *IET Image Processing*, *17*(1), 227–238. https://doi.org/10.1049/ipr2.12630

[8] Johnson, F. T., Joseph, E. O., & Folasade, A. A. (2024). Security concerns in electronic files authenticated systems. *COJ Robotics & Artificial Intelligence*, *3*(3), 000565.

[9] Ammour, B., Boubchir, L., Bouden, T., & Ramdani, M. (2020). Face–iris multimodal biometric identification system. *Electronics*, *9*(1), 85. https://doi.org/10.3390/electronics9010085

[10] Johnson, F., Oluwatobi, O., Folorunso, O., Ojumu, A. V., & Quadri, A. (2023). Optimized ensemble machine learning model for software bugs prediction. *Innovations in Systems and Software Engineering*, *19*(1), 91–101. https://doi.org/10.1007/s11334-022-00506-x

[11] Szymkowski, M., Jasiński, P., & Saeed, K. (2021). Iris-based human identity recognition with machine learning methods and discrete fast Fourier transform. *Innovations in Systems and Software Engineering*, *17*(3), 309–317. https://doi.org/10.1007/s11334-021-00392-9

[12] Kanade, S., Petrovska-Delacrétaz, D., & Dorizzi, B. (2010). Obtaining cryptographic keys using feature level fusion of iris and face biometrics for secure user authentication. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, 138–145. https://doi.org/10.1109/CVPRW.2010.5544618

[13] Joseph, E. O., Temitope, J. F., & Folasade, A. A. (2024). SEC-SKETCH: A secret-sketch graphical authentication system. *Journal of Current Trends in Computer Science Research*, *3*(5), 1–9.

[14] Ibtehaz, N., Chowdhury, M. E. H., Khandakar, A., Kiranyaz, S., Rahman, M. S., Tahir, A., . . . , & Rahman, T. (2022). EDITH: ECG biometrics aided by deep learning for reliable individual authentication. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *6*(4), 928–940. https://doi.org/10.1109/TETCI.2021.3131374

[15] Raja, K. B., Raghavendra, R., Stokkenes, M., & Busch, C. (2015). Multi-modal authentication system for smartphones using face, iris and periocular. In *International Conference on Biometrics*, 143–150. https://doi.org/10.1109/ICB.2015.7139044

[16] Yang, W., Wang, S., Shahzad, M., & Zhou, W. (2021). A cancelable biometric authentication system based on feature-adaptive random projection. *Journal of Information Security and Applications*, *58*, 102704. https://doi.org/10.1016/j.jisa.2020.102704

[17] Yang, W., Wang, S., Hu, J., Ibrahim, A., Zheng, G., Macedo, M. J., . . . , & Valli, C. (2019). A cancelable iris-and steganography-based user authentication system for the internet of things. *Sensors*, *19*(13), 2985. https://doi.org/10.3390/s19132985

[18] Baha, E., Fadhel, A., Buenaventura, P., Yeun, C. Y., Zemerly, J., & Eldelbi, K. (2024). Multimodal biometric authentication systems: Exploring iris and EEG data. In *2nd International Conference on Cyber Resilience*, 1–4. https://doi.org/10.1109/ICCR61006.2024.10533134

[19] Oren, M., Papageorgiou, C., Sinha, P., Osuna, E., & Poggio, T. (1997). Pedestrian detection using wavelet templates. In *Proceedings of IEEE Computer Society Conference on*

Computer Vision and Pattern Recognition, 193–199. https://doi.org/10.1109/CVPR.1997.609319

[20] Ali, H. H., & Shaker, S. H. (2023). Secured E-voting system based on iris identification. In *AIP Conference Proceedings*, *2591*(1), 030025. https://doi.org/10.1063/5.0119826

[21] Mohammed, A. H. Y., Dziyauddin, R. A., Kamaruddin, N., & Rahim, F. A. (2025). A hybrid ranking algorithm for secure and efficient iris template protection. *Computers & Security*, *150*, 104216. https://doi.org/10.1016/j.cose.2024.104216

[22] Singla, D., & Verma, N. (2024). Performance analysis of authentication system: A systematic literature review. *Recent Advances in Computer Science and Communications*, *17*(7), 47–67. https://doi.org/10.2174/0126662558246531231121115514

[23] Sandhya, M., Morampudi, M. K., Pruthweraaj, I., & Garepally, P. S. (2023). Multi-instance cancelable iris authentication system using triplet loss for deep learning models. *The Visual Computer*, *39*(4), 1571–1581. https://doi.org/10.1007/s00371-022-02429-x

[24] Hosseinzadeh, M., Vo, B., Ghafour, M. Y., & Naghipour, S. (2021). Electrocardiogram signals-based user authentication systems using soft computing techniques. *Artificial Intelligence Review*, *54*, 667–709. https://doi.org/10.1007/s10462-020-09863-0

[25] Sajjad, M., Khan, S., Hussain, T., Muhammad, K., Sangaiah, A. K., Castiglione, A., . . . , & Baik, S. W. (2019). CNN-based anti-spoofing two-tier multi-factor authentication system. *Pattern Recognition Letters*, *126*, 123–131. https://doi.org/10.1016/j.patrec.2018.02.015

[26] Ogbanufe, O., & Kim, D. J. (2018). Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment. *Decision Support Systems*, *106*, 1–14. https://doi.org/10.1016/j.dss.2017.11.003

[27] Vazquez-Fernandez, E., & Gonzalez-Jimenez, D. (2016). Face recognition for authentication on mobile devices. *Image and Vision Computing*, *55*, 31–33. https://doi.org/10.1016/j.imavis.2016.03.018

[28] Ackerson, J. M., Dave, R., & Seliya, N. (2021). Applications of recurrent neural network for biometric authentication & anomaly detection. *Information*, *12*(7), 272. https://doi.org/10.3390/info12070272

[29] Kim, J. J., & Hong, S. P. (2016). Design of a secure biometric authentication framework using PKI and FIDO in fintech environments. *International Journal of Security and Its Applications*, *10*(12), 69–80. https://doi.org/10.14257/ijsia.2016.10.12.07

[30] Ibrahim, A., & Ouda, A. (2017). A hybrid-based filtering approach for user authentication. In *IEEE 30th Canadian Conference on Electrical and Computer Engineering*, 1–5. https://doi.org/10.1109/CCECE.2017.7946830

[31] Ying, B., & Nayak, A. (2019). Lightweight remote user authentication protocol for multi-server 5G networks using self-certified public key cryptography. *Journal of Network and Computer Applications*, *131*, 66–74. https://doi.org/10.1016/j.jnca.2019.01.017

[32] Punithavathi, P., Geetha, S., Karuppiah, M., Islam, S. H., Hassan, M. M., & Choo, K. K. R. (2019). A lightweight machine learning-based authentication framework for smart IoT devices. *Information Sciences*, *484*, 255–268. https://doi.org/10.1016/j.ins.2019.01.073

[33] El-Sofany, H., Bouallegue, B., & Abd El-Latif, Y. M. (2024). A proposed biometric authentication hybrid approach using iris recognition for improving cloud security. *Heliyon*, *10*(16), e36390. https://doi.org/10.1016/j.heliyon.2024.e36390

[34] Adamović, S., Miškovic, V., Maček, N., Milosavljević, M., Šarac, M., Saračević, M., & Gnjatović, M. (2020). An efficient novel approach for iris recognition based on stylometric features and machine learning techniques. *Future Generation Computer Systems*, *107*, 144–157. https://doi.org/10.1016/j.future.2020.01.056

[35] Tahir, A. A., & Anghelus, S. (2022). Improving iris recognition accuracy using Gabor kernels with near-horizontal orientations. *International Journal of Advances in Signal and Image Sciences*, *8*(1), 25–39. https://doi.org/10.29284/ijasis.8.1.2022.25-39

[36] Hafeez, H., Zafar, M. N., Abbas, C. A., Elahi, H., & Ali, M. O. (2022). Real-time human authentication system based on iris recognition. *Eng*, *3*(4), 693–708. https://doi.org/10.3390/eng3040047

[37] Ullah, A., Salam, A., El-Raoui, H., Sebai, D., & Rafie, M. (2022). Towards more accurate iris recognition system by using hybrid approach for feature extraction along with classifier. *International Journal of Reconfigurable and Embedded Systems*, *11*(1), 59–70. https://doi.org/10.11591/ijres.v11.i1.pp59-70

[38] Moi, S. H., Yong, P. Y., Hassan, R., Asmuni, H., Mohamad, R., Weng, F. C., & Kasim, S. (2022). An improved approach of iris biometric authentication performance and security with cryptography and error correction codes. *JOIV: International Journal on Informatics Visualization*, *6*(2–2), 531–539. https://doi.org/10.30630/joiv.6.2-2.1091

How to Cite: Ajayi, F., Aromolaran, A., Victor, A., Usman, A.-A., & Johnson, F. (2025). EN-IRAM: Enhance Iris Authentication Model with Soft-Vote Ensemble Learner and Gradual Feature Selector. *Artificial Intelligence and Applications*, *3*(3), 295–304. https://doi.org/10.47852/bonviewAIA52023875