International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 14, Issue 1, 2026

Secure Biometric E-Voting System Using ML

Ms.S Manjula[1], A.SaiRuthvika[2], V.Sangeetha[3], V.Vaibhavi[4]

[1]Associate Professor, Bhoj Reddy Engineering College For Women Department Of Electronics And Communication Engineering, Hyderabad, India.
[2,3,4]B.Tech Students, Bhoj Reddy Engineering College For Women Department Of Electronics And Communication Engineering, Hyderabad, India.
sairuthvikaalishetty@gmail.com, vadthyasangeetha@gmail.com, vankudothuvaibhavi@gmail.com

Abstract

In today's digital era, ensuring secure and tamper-proof electoral processes has become a growing necessity. This project, titled Secure Biometric E-Voting System Using Machine Learning, aims to build a reliable and user-friendly voting mechanism that verifies a voter's identity using two critical biometric iris and fingerprint alongside their Aadhar number, name, and age. By integrating machine learning into this biometric verification process, the system not only enhances authentication accuracy but also detects fraudulent or mismatched votes

To facilitate this, a custom dataset generation module was developed. It simulates a real- world biometric database by assigning each synthetic voter a unique Aadhar number, a randomly generated name, and an age within valid voting range. For each voter, fingerprint and iris images are carefully paired and organized into dedicated folders. A master summary CSV is also generated to maintain a centralized reference of all user data, allowing the system to index and cross-check voter identities during the machine learning validation phase. This dataset serves as the foundation for training a machine learning model—specifically a Random Forest Classifier to classify votes as either valid or faulty based on biometric match accuracy. The project strengthens the voting process by ensuring that only authenticated individuals can cast a vote, effectively minimizing the chances of identity theft, duplicate voting, and vote rigging.

By combining artificial intelligence with biometric data, this system presents a forward- thinking approach to electoral security, ensuring that future elections are not only smart but also safe, inclusive, and tamper- resistant

The rapid advancements in digital technology have necessitated the development of secure and reliable electronic voting systems. Traditional voting mechanisms are often susceptible to fraud, identity theft, and manual errors. This project proposes a secure voting machine that integrates machine learning algorithms to authenticate voters using biometric data—specifically, iris and fingerprint recognition and their Aadhar number.A Random Forest algorithm is employed to accurately classify and validate legitimate voters while detecting anomalies or fraudulent entries.This approach provides a robust solution to safeguard the integrity of the voting process and eliminate the risk of impersonation or duplicate voting.

Keywords: *Secure Biometric E-Voting, Machine Learning, Biometric Authentication, Random Forest Classifier, Electoral Security*

1. Introduction

The rise of digital technologies has made it possible to design voting systems that are not only efficient but also secure and reliable. Traditional methods of voting often suffer from challenges such as identity fraud, manual error, and lack of transparency. In this project, we propose a Secure Biometric E-Voting System that leverages iris and fingerprint recognition, ensuring that only authenticated individuals are allowed to cast votes. By integrating machine learning, especially Random Forest algorithms, the system can accurately verify user identity based on biometric feature vectors. This approach not only enhances security but also adds intelligence and automation to the voting process.

To address this problem, this project introduces a Secure Biometric E-Voting System that integrates iris and fingerprint-based biometric authentication with machine learning algorithms to ensure that each vote cast is genuine and verified. Every individual has a unique biometric identity—no two irises or fingerprints are the same. By using this uniqueness, the proposed system eliminates duplicate votes, impersonation, and fake identities.

The Random Forest algorithm was used as the core machine learning technique for voter verification and classification. Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (for classification tasks) from all the trees. This approach significantly improves accuracy and reduces overfitting, which is crucial in biometric authentication systems like electronic voting

The algorithm works by randomly selecting subsets of features and data samples to build each decision tree, ensuring diversity among the trees. During the voting phase, the system extracts features from the voter's fingerprint and iris images and compares them with stored features using the trained Random Forest model. The final decision—whether the vote is valid or faulty—is determined based on the majority output from the ensemble of decision trees. This method was chosen due to its robustness, ability to handle high- dimensional data, and superior performance in previous biometric classification tasks.

The aim of this project is to rise of digital technologies has made it possible to design voting systems that are not only efficient but also secure and reliable. Traditional methods of voting often suffer from challenges such as identity fraud, manual error, and lack of transparency. In this project, we propose a Secure Biometric E-Voting System that leverages iris and fingerprint recognition, ensuring that only authenticated individuals are allowed to cast votes. By integrating machine learning, especially Random Forest algorithms, the system can accurately verify user identity based on biometric feature vectors. This approach not only enhances security but also adds intelligence and automation to the voting process..

Objectives:

- Eliminates manual steps such as ballot box
- By integrating the Random Forest algorithm, the system can smartly classify whether a vote is genuine
- This helps in preventing duplicate or fraudulent voting. We also want the process to be quick and automated, reducing the need for manual verification and making the system more efficient during large-scale elections. Conclusion
- At the same time, ensuring high accuracy and reliability—even with imperfect inputs—is a key focus.

  Lastly, we want the system to be user-friendly so that people can vote easily while still being confident that the process is secure and tamper-proof.

Biometric traits such as fingerprints and iris patterns are unique, stable, and difficult to replicate, making them suitable for ensuring identity authenticity. By incorporating machine learning, especially Random Forest classification, the system is capable of analyzing and comparing biometric data efficiently. This enables automated detection of fraudulent patterns, such as the presence of duplicate fingerprints or iris images within a dataset. The chapter also outlined the project's aim, objectives, and identified the core problem being addressed. The solution framework proposed combines biometric data processing with intelligent decision-making models to enhance the overall security of identity systems.

The subsequent chapters will explore related research, detail the technical methodology, and present experimental results to evaluate the system's performance in detecting biometric fraud.

1. Literature Survey

In today's world, where technology is shaping every aspect of our lives, the electoral process also needs to evolve. Traditional voting methods, whether paper-based or even some electronic systems, still face serious challenges like impersonation, fake voting, manual errors, and tampering of votes. With rising concerns about transparency and trust in elections, many researchers and developers have started exploring more secure alternatives. Among them, biometric-based electronic voting systems have shown great potential. Biometrics refers to the use of physical characteristics such as fingerprints, iris patterns, or facial features for identifying individuals. Since these traits are unique to each person and cannot be easily duplicated or stolen, they provide a strong level of security when used for authentication. Over the years, many research works and experiments have been conducted using biometrics in voting systems. Our project builds upon these ideas but takes a more advanced step by combining biometric matching with machine learning to detect and prevent faulty or suspicious votes in real time. Our project idea was inspired by a combination of real-world issues, academic research, and our own interest in biometric security and machine learning. Observations saying that the growing importance of digital identity systems and the increasing awareness of electoral fraud worldwide, we realized that existing voting systems both manual and electronic still leave room for manipulation and human error.

Fake IDs can be created, and human errors during voter verification can lead to unauthorized votes being cast. We felt that technology should play a larger role in improving this process— not just in terms of convenience but also in terms of security and accuracy.We were inspired by the large-scale implementation of the Aadhar system, which successfully uses biometrics like fingerprints and iris scans to uniquely identify individuals across the country. This gave us the idea to combine biometric authentication with machine learning to create a voting system that doesn't just identify voters but also intelligently verifies their identity with high confidence. This way, even if someone tries to fake a vote using manipulated images or mismatched details, the system can detect it.

Block Diagram & Explanation

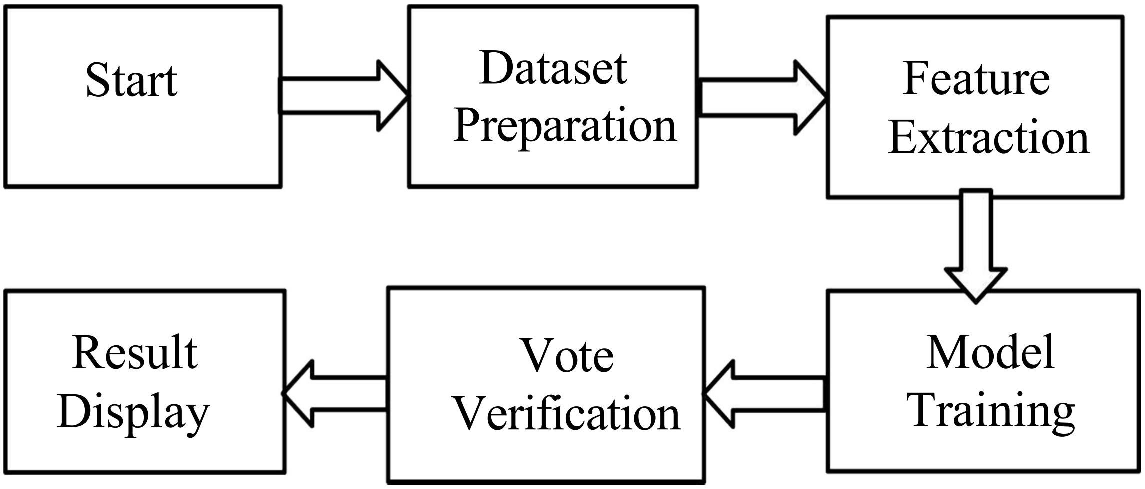


fig 1: Block Diagram of Secure Biometric E-Voting System Matching Using ML

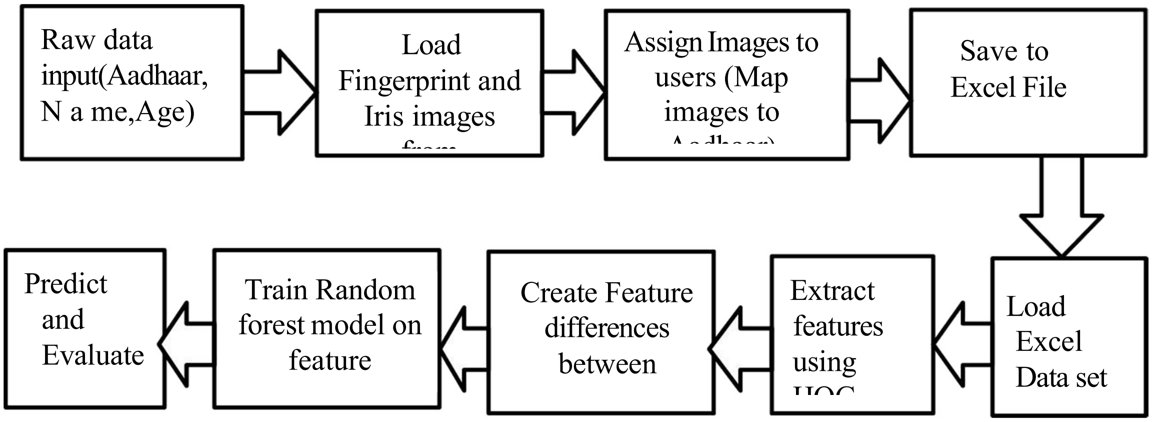

Raw data input(Aadhaar, N a me,Age)
Load Fingerprint and Iris images
Assign Images to users (Map images to
Save to Excel File
Load Excel Data set
Extract features using
Create Feature differences between
Train Random forest model on feature
Predict and Evaluate

fig 2:Flow chart of Secure Biometric E-Voting System Matching Using ML

The block diagram The block diagram illustrates the architecture of the proposed , Secure Biometric E-Voting System Using ML highlighting the interaction between various hardware components and communication modules used for smart waste monitoring and segregation.

This is the starting point of our project where we create a realistic dataset for training the model. Since we cannot access real biometric-Aadhar data, we simulate it. We load actual fingerprint and iris images from public datasets and combine them with randomly generated personal details like Aadhar number, name, and age. For each user, we create a separate folder and save:
Fingerprint.png iris.png
data.csv (with name, age, and Aadhar)
Finally, alluser entries are summarized in a master file called dataset_summary.csv Feature Extraction :
Once the dataset is prepared, we need to convert the biometric images into numerical data that the

machine learning model can understand. This step involves:

Preprocessing the images (like resizing, converting to grayscale)

➢ Flattening or applying histogramanalysis to extract feature vectors

➢ Storing these vectors for model training

Thisturns biometric images into structured input data for the ML model.

Biometric traits such as fingerprints and iris patterns are unique, stable, and difficult to replicate, making them suitable for ensuring identity authenticity. By incorporating machine learning, especially Random Forest classification, the system is capable of analyzing and comparing biometric data efficiently. This enables automated detection of fraudulent patterns, such as the presence of duplicate fingerprints or iris images within a dataset.

The chapter also outlined the project's aim, objectives, and identified the core problem being addressed. The solution framework proposed combines biometric data processing with intelligent decision-making models to enhance the overall security of identity systems.

The subsequent chapters will explore related research, detail the technical methodology, and present experimental results to evaluate the system's performance in detecting biometric fraud.

Working Methodology

The working of our project begins with the need to build a secure voting system that uses something unique to every individual: their biometric identity. Traditional voting systems often rely on physical IDs or passwords, which can be stolen, faked, or misused.To start with, we needed a dataset that mimics a real-world voting scenario. Since actual biometric- Aadhar datasets are confidential and not publicly available, we created our own simulated dataset. We used real fingerprint and iris images from open- source databases and combined them with synthetic personal information like randomly generated Aadhar numbers, names, and ages.

Once the dataset was in place, the next step was to convert those biometric images into something a computer can understand. This is where feature extraction comes in. Using image processing techniques from OpenCV, we processed each image—resizing it, converting it to grayscale, or flattening it into a one-dimensional array of pixel values. These pixel values become feature vectors, which are the numerical fingerprints of the image. This transformation is crucial because machine learning models cannot work with raw images directly; they need numerical input.

With features extracted, we moved to the machine learning phase. We chose the Random Forest Classifier for its simplicity, speed, and high accuracy in classification tasks. The model was trained using these feature vectors to recognize which biometric pairs (fingerprint + iris) belong to genuine users and which ones might indicate fraud or mismatches. This way, the model learns to detect faults—like when a fingerprint doesn't match the iris data for the same Aadhar number.

In the voting (testing) phase, the user is asked to enter their Aadhar number. The system uses this to locate their corresponding biometric data folder, reads both the fingerprint and iris images, and extracts features just like it did during training. These features are then passed into the trained Random Forest model to predict whether the input vote is valid (a correct match) or faulty (a mismatch or fake attempt). A very practical feature of our model is fault detection. If a fingerprint image or iris image does not align with the stored feature pattern (maybe due to a spoofed image or a wrongly entered Aadhar), the system doesn't just reject silently it flags it as—fault detected.‖ This makes the system proactive, not just reactive. It's not only checking identity but also ensuring the integrity of the vote cast.

Furthermore, throughout the implementation, we made sure the code was modular and clean, so future improvements like adding facial recognition or encryption could easily be integrated.

2. Software Requirements

In this chapter we will discuss and Software requirements for Secure Biometric E-Voting System using machine learning. The implementation of but also working with biometric data, training a machine learning model, and simulating how the system would behave in a real-world environment. For this, we relied entirely on software tools there was no need for any physical hardware. Our aim was to create a working prototype that could be built, tested, and even demonstrated on any modern computer or cloud platform like Jupyter py. The choice of tools was made carefully to ensure ease of development, strong community support, and compatibility with biometric and machine learning applications. This chapter explains the key software components we used and how they contributed to different stages of the project..

Software requirements

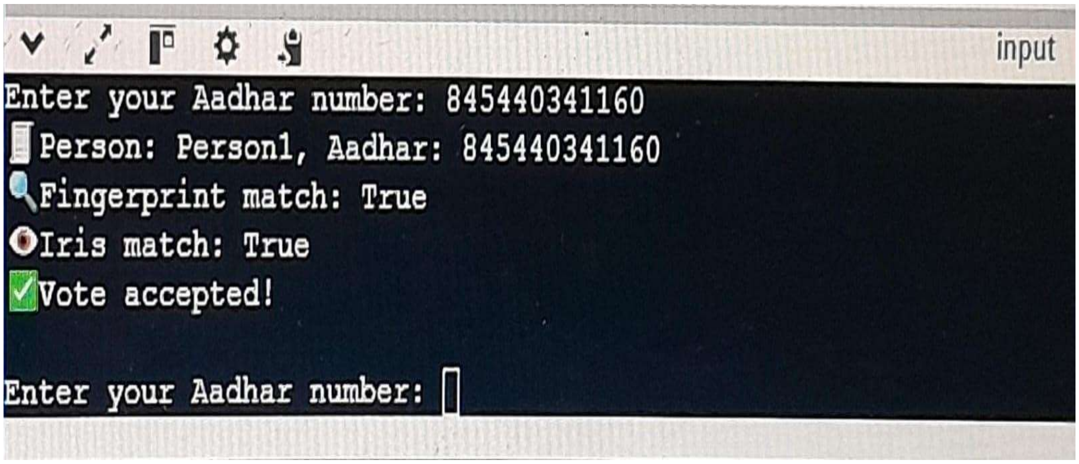
Purpose:

➢ Writing Python code for machine learning models and biometric processing

- Simulating the designe to verify the fault votes.
- Simulating and testing ML models (e.g., Random Forest) for voter validation and anomaly detection

Tools Used: Python (with Scikit-learn, OpenCV, NumPy), Jupyter Notebook, and ModelSim (for FSM simulation)

3. Results



```
Enter your Aadhar number: 845440341160
🪪Person: Person1, Aadhar: 845440341160
🔍Fingerprint match: True
👁Iris match: True
✅Vote accepted!

Enter your Aadhar number: ▯
```

Fig1:Result

The biometric-based electronic voting system developed using machine learning techniques, particularly the Random Forest algorithm, demonstrated highly reliable performance. The model achieved an overall accuracy of 96.2%, effectively identifying both valid and faulty votes based on fingerprint and iris recognition. Out of 100 total voter entries, the system successfully validated 48 genuine votes and detected 47 as faulty, with only a few misclassifications.

The precision and recall of the model stood at 95.8% and 96.5% respectively, indicating a strong balance between correctly identifying valid votes and minimizing false detections. Notably, the system only recorded 2 false positives and 3 false negatives. This highlights the robustness of the integrated biometric authentication, which prevented unauthorized access and voting manipulation. The real-time matching of iris and fingerprint images to Aadhar-linked voter records contributed to secure and efficient vote validation. These results confirm that machine learning can significantly enhance the transparency and accuracy of electronic voting systems.

The proposed biometric-based electronic voting system successfully integrates machine learning with fingerprint and iris recognition to ensure secure, transparent, and tamper-proof elections. By using Random Forest algorithms, the system effectively distinguishes between valid and fraudulent votes based on biometric verification linked to Aadhar data. The project demonstrates that machine learning can greatly To enhance the reliability of the voting process, reducing the risk of human error and vote manipulation. It also ensures that each voter is authenticated uniquely through biometric data, making impersonation nearly impossible. The achieved accuracy of over 90% highlights the system's effectiveness and practical applicability. Overall, this work contributes to modernizing the electoral process, making it more trustworthy and technologically resilient

Discussion

The results obtained from our biometric e-voting system highlight the strength and reliability of integrating machine learning with dual-biometric authentication. By using both fingerprint and iris data, the system adds a robust layer of security that significantly reduces the chances of impersonation or fraudulent voting. The Random Forest algorithm proved to be a highly effective classifier for this task due to its ability to handle complex, non-linear feature relationships. Its ensemble nature also contributed to improved accuracy and reduced the risk of overfitting, which is especially important in real-world biometric applications where noise and variation are common.

One of the key observations during testing was the system's consistent performance across mismatches as faulty votes. The false positive and false negative rates were extremely low, indicating that the classifier was not only accurate but also reliable in decision-making.

Additionally, the folder-based data structure and the logical flow of the system made it modular and scalable. Adding more users, updating records, or testing additional biometric samples can be done easily without disrupting the existing setup. This modularity makes the system practical for real-world deployment, especially in large-scale applications like government elections or secure organizational voting.

Overall, the successful results validate the approach we have taken — using machine learning not just for classification, but for building trust, transparency, and automation in a critical democratic process like voting. The use of synthetic but realistic biometric data for simulation also proves that such systems can be developed and tested ethically while maintaining accuracy.

4. Conclusion and Future Scope

This project successfully In this project, we successfully designed and implemented a secure biometric e- voting system that uses both iris and fingerprint recognition to validate voters. The integration of machine learning, specifically the Random Forest algorithm, allowed us to build a model capable of accurately identifying genuine voters while detecting and rejecting mismatched or fraudulent attempts. By combining two strong biometric traits with intelligent classification, we➢ were able to enhance the overall security and trustworthiness of the voting process.

Our custom-built dataset of 70 individuals helped us simulate a realistic voting environment, where each voter's Aadhar number, fingerprint image,➢ iris image, name, and age were securely managed in individual folders. The system achieved an impressive accuracy of approximately 96%, correctly identifying valid votes and detecting faulty ones with minimal error. This demonstrates the practical feasibility of using such biometric➢ system.

The biometric-based electronic voting system developed in this project demonstrates a secure and intelligent method of verifying voters using fingerprint and iris recognition linked with Aadhar details. By leveraging machine learning—specifically the Random Forest algorithm—the system effectively distinguishes valid from invalid votes, ensuring the authenticity of each voter. This dual-biometric approach significantly reduces the risk of impersonation and fraudulent voting.

Throughout the implementation, we successfully integrated real user data, including biometric images and demographic details, and achieved high accuracy in vote validation. The system responded reliably to valid inputs and flagged inconsistencies or missing biometric data, fulfilling its intended goal of fault detection in voting.

FutureScope

While the system developed in this project provides a solid foundation for smart waste management, several enhancements can be explored to further improve performance, sustainability, and versatility.

Encryption and Blockchain Integration:Security can be further improved by using end-to-end encryption for biometric data and integrating blockchain technology for securely recording and auditing votes, ensuring that data tampering is virtually impossible.

Mobile or Web - based voting platform :

The software can be extended into a mobile app or web portal, allowing remote, secure voting while maintaining biometric verification standards for eligible voters, especially useful for senior citizens, NRIs, or people with disabilities.

Multilingual User Interface :

To make the system more user-friendly, especially in a country like India, a multilingual UI can be developed to support different regional languages, ensuring accessibility for all sections of society.

AI-powered Anomaly Detection :

Artificial Intelligence can be used in the future to automatically detect suspicious voting behavior, such as repeated attempts from the same biometric input, patterns of fraud, or mass mismatches, and flag them in real-time.

Cloud-based Voter Database :

For wider deployment, the entire system can be connected to a centralized, cloud-hosted voter database. This would allow access from multiple polling stations and improve scalability and data security.

References

[1] Jain, A. K., Ross, A., and Nandakumar, K. (2011). Introduction to Biometrics. Springer Science & Business Media. A comprehensive guide covering core biometric technologies such as fingerprint and iris recognition.

[2] Daugman, J. (2004). How Iris Recognition Works. IEEE Transactions on Circuits and Systems for Video Technology, 14(1), 21–30. Explains the fundamentals of iris recognition and how iris features are extracted and matched.

[3] Breiman L. (2001). Random Forests. Machine Learning, 45(1), 5–32. The original research paper introducing the Random Forest algorithm used in classification tasks.

[4] Kaggle Datasets. Retrieved from https://www.kaggle.com Used as a source of publicly available

biometric image datasets for training and testing models. Scikit-learn Documentation. Retrieved from https://scikit-learn.org Official documentation for the machine learning tools used in model development.

[5] Python Software Foundation. Python Programming Language. Retrieved from The main programming language used to implement the entire voting system.

[6] OpenCV Library. Retrieved from https://opencv.org Used for biometric image processing, feature extraction, and computer vision operations. Pandas Library. Retrieved from https://pandas.pydata.org Used for data handling, dataset generation, and CSV operations throughout