Privacy Preserving Face Recognition Attendance System

Major Project Report

Submitted for the partial fulfillment of the degree of

Bachelor of Technology

In

Computer Science & Design

Submitted By

Gaurav Sharma

0901CD211026

UNDER THE SUPERVISION AND GUIDANCE OF

Dr. Rohit Agrawal

Assistant Professor

Department of Computer Science & Engineering

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR (M.P.), INDIA
माधव प्रौद्योगिकी एवं विज्ञान संस्थान, ग्वालियर (म.प्र.), भारत
(Deemed to be University)
NAAC ACCREDITED WITH A++ GRADE

January-May 2025

DECLARATION BY THE CANDIDATE

I hereby declare that the work entitled "Privacy Preserving Face Recognition Attendance System" is my work, conducted under the supervision of **Dr. Rohit Agrawal,** during the session Jan-May 2025. The report submitted by me is a record of bonafide work carried out by me.

I further declare that the work reported in this report has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Gaurav Sharma

0901CD211026

Date: 21/5/25
Place: Gwalior

This is to certify that the above statement made by the candidates is correct to the best of my knowledge and belief.

Guided By:

21/05/2025

Dr. Rohit Agrawal
Assistant Professor
Computer Science and Engineering
MITS, Gwalior

Departmental Project Coordinator

21/05/25

Dr. Kuldeep Narayan Tripathi
Assistant Professor
Computer Science and Engineering
MITS, Gwalior

Approved by Head

21/05/25

Dr. Manish Dixit

Dr. Manish Dixit
Computer Science and Engineering
MITS, Gwalior

PLAGIARISM CHECK CERTIFICATE

This is to certify that I, a student of B.Tech. in **Computer Science & Engineering** have checked my complete report entitled "Privacy Preserving Face Recognition Attendance System" for similarity/plagiarism using the "Turnitin" software available in the institute.

This is to certify that the similarity in my report is found to be ...15%. which is within the specified limit (30%).

The full plagiarism report along with the summary is enclosed.

Gaurav

Gaurav Sharma

0901CD211026

Checked & Approved By:

Prof. Mahesh Parmar
Assistant Professor
Computer Science and Engineering
MITS, Gwalior

ABSTRACT

In recent years, face recognition technology has gained widespread adoption for automated attendance systems due to its non-intrusive and efficient nature. However, conventional face recognition methods raise significant privacy concerns, especially when sensitive biometric data is stored and transmitted in unencrypted formats. This project aims to address these concerns by implementing a privacy-preserving face recognition attendance system that ensures data security while maintaining system accuracy and usability.

The system captures facial images of individuals using a camera, processes them locally to extract facial features, and encrypts this biometric data before storage or transmission. Techniques such as AES encryption are applied to protect face encodings, ensuring that raw facial data is never exposed. Furthermore, a secure authentication and decryption mechanism is integrated to allow only authorized personnel to access or manage the data. The backend is implemented using a secure database (MongoDB), while the front end provides a user-friendly interface for attendance tracking and management.

This project not only demonstrates the practical feasibility of combining face recognition with cryptographic techniques but also contributes to building trust in biometric systems. By safeguarding sensitive data, it mitigates the risk of identity theft and unauthorized surveillance, thus making it suitable for educational institutions and workplaces that prioritize data privacy. The approach lays the groundwork for more robust and ethically responsible biometric systems in the future.

ACKNOWLEDGEMENT

The whole project on its own has proved to be extremely helpful to my career. I am thankful to my institute, Madhav Institute of Technology and Science, for allowing me to continue my disciplinary/interdisciplinary project as a curriculum requirement under the provisions of the Flexible Curriculum Scheme (based on the AICTE Model Curriculum 2018), approved by the Academic Council of the institute. I extend my gratitude to the Vice Chancellor of the institute, **Dr R.K. Pandit** and the Dean, Faculty of Engineering & Technology, **Dr Manjaree Pandit**, for this opportunity.

I would sincerely like to thank my department, the Department of Computer Science and Engineering, for allowing me to explore this project. I humbly thank **Dr. Manish Dixit, Professor** and Head of the Department of Computer Science and Engineering, for his continued support during this engagement, which eased the process and formalities involved. Additionally, I am sincerely thankful to my faculty mentor, **Dr. Gagandeep Kaur**, Assistant Professor, Department of Computer Science and Engineering, for her continued support and guidance throughout the project. I am also very thankful to the department's faculty and staff.

Gaurav Sharma

0901CD211026

CONTENT

Table of Contents

CHAPTER 1: INTRODUCTION

Face recognition, being contactless and fast, offers a modern solution to these challenges by automating attendance processes efficiently. However, with growing concerns over personal data security and privacy, especially in biometric systems, it is crucial to develop solutions that not only enhance convenience but also safeguard user identities. This project addresses that need by developing a privacy-preserving face recognition attendance system.

In conventional face recognition systems, facial data is typically stored in raw or easily reversible formats, making it vulnerable to data breaches and unauthorized access. Given that facial data is a unique and permanent identifier, its misuse can lead to severe consequences, including identity theft and surveillance abuse. Recognizing these risks, our system focuses on incorporating privacy-preserving mechanisms, such as data encryption, secure storage, and controlled access. By integrating cryptographic techniques—like AES encryption for facial embeddings—and ensuring that data is never stored or transmitted in plain form, the system ensures a high level of data confidentiality and integrity. Only authorized users with decryption keys can access or manage the biometric data, reducing the risk of unauthorized exploitation.

The project combines several technologies and methodologies to achieve its goal. We use OpenCV and deep learning models for real-time face detection and recognition, ensuring high accuracy and speed. Python is used as the primary programming language, with modules such as cv2, face_recognition, and cryptography for image processing and secure data handling. MongoDB serves as the backend database, offering a flexible and scalable solution for storing encrypted attendance records. The front end provides a simple and intuitive interface for managing attendance logs and verifying users. Overall, this system not only demonstrates the practical utility of integrating machine learning with data security but also sets a benchmark for ethically responsible use of biometrics in daily operations.

CHAPTER 2: LITERATURE SURVEY / TECHNOLOGY

Literature Survey

The concept of automated attendance using face recognition has been explored extensively in recent academic and industrial research. Traditional attendance systems—like manual registers, RFID tags, or biometric fingerprints—have shown limitations in terms of efficiency, accuracy, and hygiene. Face recognition technology has emerged as a superior alternative due to its contactless nature and ease of use. Researchers have employed various algorithms, such as Eigenfaces, Fisherfaces, and Local Binary Patterns Histogram (LBPH), to improve recognition accuracy under different lighting and pose conditions. However, most of these systems do not prioritize user privacy, and biometric data is often stored in raw form, exposing individuals to risks of identity theft. Recent studies have advocated for integrating cryptographic techniques with biometric systems to ensure data privacy and compliance with data protection laws like GDPR. This project builds upon such research by incorporating encryption with LBPH-based face recognition.

Technologies Used

The project is developed using Python as the core programming language due to its rich ecosystem of libraries and frameworks for image processing, GUI development, and database connectivity. OpenCV is used for face detection and recognition, employing the LBPH algorithm, which is particularly effective in handling varying lighting conditions. The tkinter library powers the graphical user interface, offering a user-friendly experience for both administrators and users. MongoDB, a NoSQL database, is used for storing encrypted facial images, making it easier to manage and retrieve unstructured data. To ensure data privacy, the cryptography library is used with Fernet symmetric encryption to encrypt and decrypt face images securely. Additional tools like dotenv are used for managing environment variables such as database URIs and encryption keys, promoting secure and scalable deployment. This technology stack together forms a robust, secure, and interactive attendance system that prioritizes both functionality and privacy.

CHAPTER 3: METHODOLOGY

This project follows a systematic approach to ensure a seamless and secure user experience for automated attendance marking using face recognition. The entire system is designed with a strong focus on privacy, leveraging encryption and secure storage.

3.1. Environment Setup and Initialization

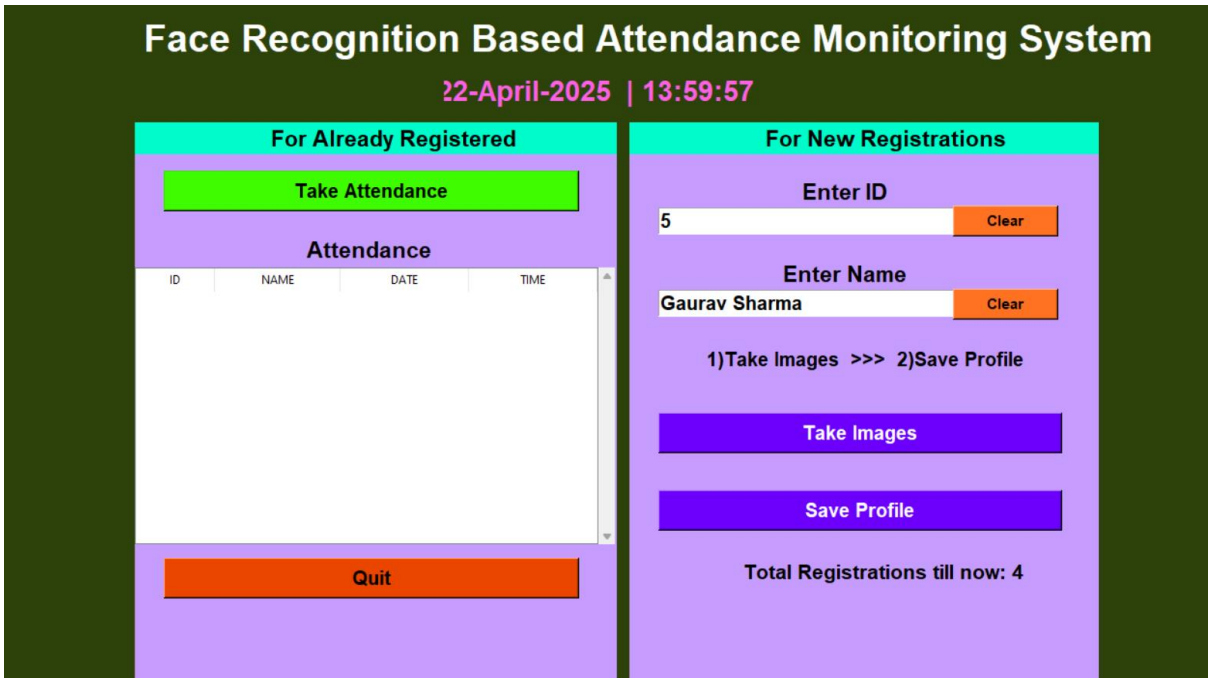
To begin with, the necessary environment is set up to support image processing, encryption, and GUI functionalities. This includes installing Python packages like OpenCV, NumPy, PIL, tkinter, pymongo, and cryptography. Environmental variables such as the MongoDB connection URI and encryption key are stored in a .env file for security. These are loaded into the system using the dotenv library, ensuring that no sensitive credentials are exposed in the codebase.

A MongoDB database named face_detection_db is used for storing encrypted facial data, and a collection named face_images is created to store individual image entries. The encryption key is processed using Fernet, a symmetric encryption method from the cryptography module, to handle secure encryption and decryption of biometric data.

3.2. User Registration and Image Capture

In this step, new users can register through a user-friendly GUI built with tkinter. They are prompted to enter their Name and ID. Upon clicking the "Take Images" button, the system activates the webcam and utilizes the Haarcascade classifier to detect faces in real-time.

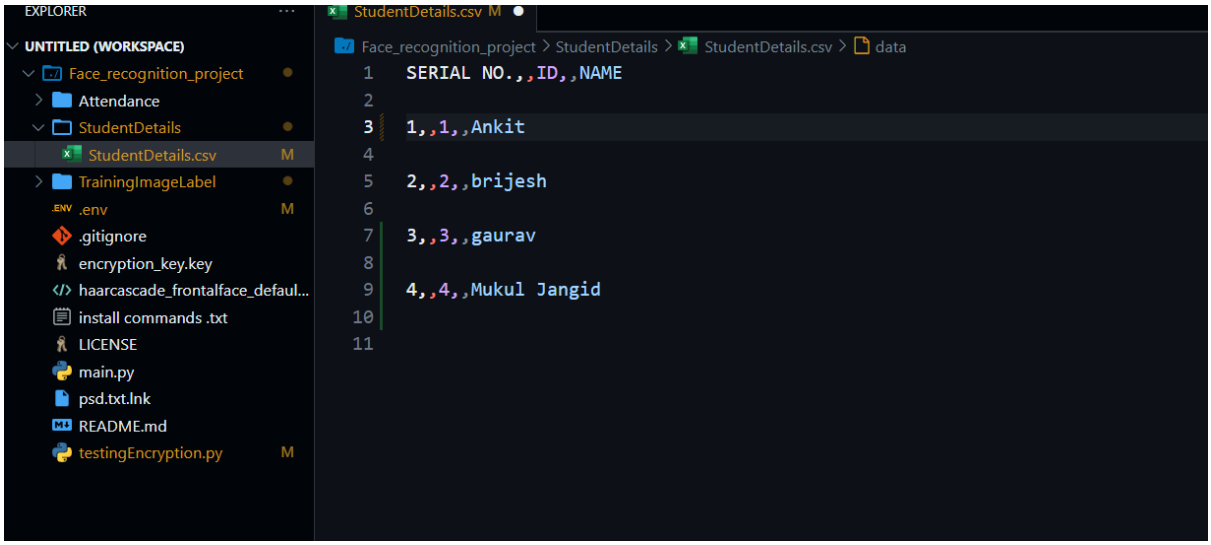
Once a face is detected, the system captures the face region, converts it to grayscale to reduce complexity, and encodes it as JPEG bytes. These bytes are then encrypted using the Fernet encryption algorithm to ensure that no raw image data is ever stored or transmitted. The encrypted images, along with metadata such as the user ID, name, and image filename, are stored in the MongoDB collection. This step ensures that biometric data is protected from unauthorized access and tampering.



3.3. Storing User Metadata

While the encrypted image data is stored securely in the database, user metadata such as ID, Name, and Serial Number is stored locally in a CSV file (StudentDetails.csv). This file is primarily used for counting the number of registered users and quickly displaying non-sensitive information in the GUI.

Using both database and file-based storage provides a balance between data security (through encrypted cloud storage) and user experience (through fast local lookups). The application ensures that every new registration is assigned a unique serial number by counting existing entries in the CSV file.



3.4. Face Recognition Model Training

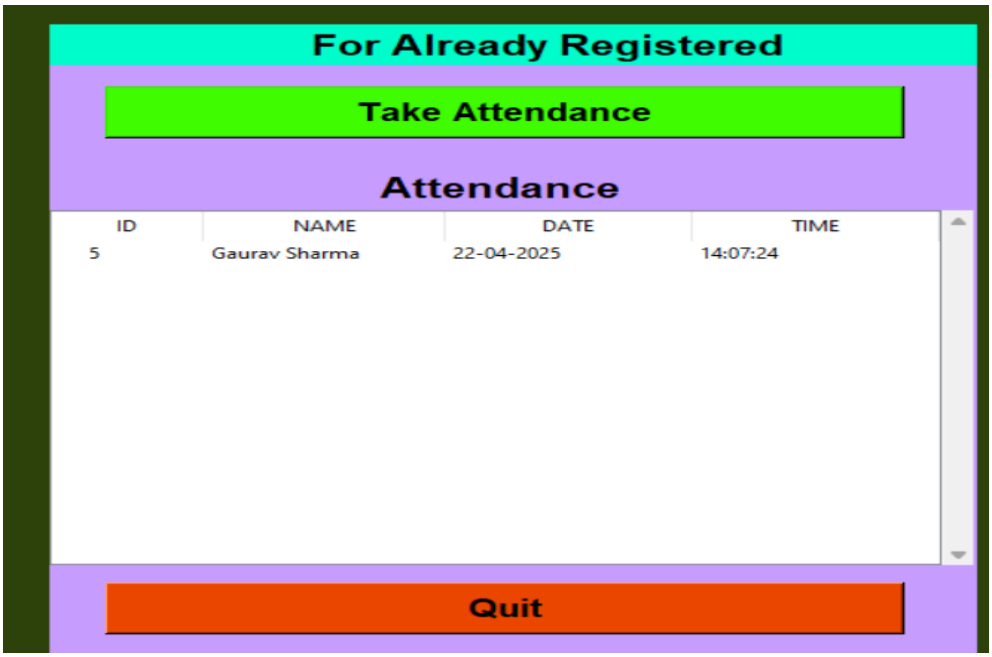
Once the facial images for a user are captured, the system allows users to save their profile by clicking the "Save Profile" button. This action is password-protected to prevent unauthorized access or retraining.

The model training involves decrypting all encrypted face images stored in MongoDB. Each image is decrypted using the same Fernet key used during encryption and then converted back into grayscale image arrays. These images, along with their respective user IDs, are then used to train an LBPH (Local Binary Patterns Histogram) face recognizer, which is particularly effective for real-time face recognition in varying lighting conditions. After training, the recognizer model is saved locally as Trainner.yml, which can be loaded later for attendance tracking.

3.5. Attendance Tracking

This step is the core functionality of the system where real-time face recognition is used to mark attendance. When the "Take Attendance" button is clicked, the trained LBPH model is loaded and the webcam is activated. Faces detected in the video stream are compared against the trained model to identify users.

If a face is recognized and the confidence score is below a specified threshold (indicating a strong match), the system records the user's ID, name, date, and time in an attendance CSV file (e.g., Attendance_22-04-2025.csv). A TreeView widget within the GUI displays the live attendance list. The system ensures that each person is marked present only once per session by keeping track of already recorded IDs in memory. This prevents duplicate entries and ensures integrity in the attendance data.



3.6. Security and Access Control

Security is a key pillar of this system. Besides encryption of facial images, password-based access control is implemented for sensitive operations like training the recognizer and changing passwords. Users are prompted to set an initial password on the first run, which is then saved locally in a secure file. Any attempt to retrain the model or change the password is restricted unless the correct password is provided.

The images used in training and recognition are never stored in raw format on disk. Only encrypted images are kept in the database, and they are decrypted only in memory during model training and face recognition. This approach significantly reduces the risk of biometric data leakage or misuse.

3.7. GUI Design and User Experience

The front end of the application is built entirely using tkinter, which offers an interactive and intuitive user interface. The layout is divided into two panels: one for new registrations and another for attendance tracking. The GUI includes clearly labeled fields and buttons, real-time clock display, date panel, and a dynamic TreeView attendance table.

Color-coded frames and messages provide visual cues for system status, such as successful registrations, errors, or attendance completion. Users are also provided with a menu bar for extra functionality, including options to change password, contact developer, and exit the application.

CHAPTER 4: CONCLUSION

The Privacy-Preserving Face Recognition Attendance System successfully demonstrates how biometric technologies can be integrated with cryptographic methods to enhance both automation and data security. By capturing facial data through a real-time camera feed, encrypting it before storage, and using secure decryption during recognition, the system ensures that sensitive information is never exposed in raw form. The use of the LBPH algorithm for face recognition, combined with MongoDB for encrypted data storage, results in a solution that is not only accurate and efficient but also highly secure. The implementation of password protection for model training and administrative access further strengthens the overall privacy model.

This project addresses the growing need for privacy-aware applications in biometric systems, especially in educational and corporate environments where data sensitivity is a concern. It lays a strong foundation for future improvements, such as incorporating advanced encryption schemes, cloud-based model training, and support for multi-camera networks. Ultimately, this work contributes to the responsible and ethical adoption of artificial intelligence in everyday use cases, highlighting that technological advancement and user privacy can go hand in hand.

REFERENCES

1. W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys (CSUR), vol. 35, no. 4, pp. 399–458, 2003.

2. L. Sweeney, "k-Anonymity: A model for protecting privacy," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 5, pp. 557–570, 2002.

3. A. Othman and A. Ross, "Privacy of facial biometrics: A survey," in Handbook of Biometrics for Forensic Science, Springer, pp. 63–78, 2014.

4. A. Pinto, F. Costa-Paz, R. Piai, and P. Escudeiro, "Privacy-preserving biometric systems: A comprehensive survey," IEEE Access, vol. 9, pp. 98824–98848, 2021.

5. OpenCV, "Face Recognition using LBPH," OpenCV Documentation, [Online]. Available: https://docs.opencv.org/4.x/dc/dc3/tutorial_py_face_recognition.html. [Accessed: Apr. 22, 2025]

TURNITIN PLAGIARISM REPORT

MPRs (If Applicable)

MONTHLY REPORT OF PROGRESS (MRP) FROM INDUSTRY MENTOR

Name of student	Gaurav Sharma		Department	CSE	
Industry/Organization	MITS		Date/Duration	15/01/2025-15/02/2025	
Criterion	**Poor**	**Average**	**Good**	**Very Good**	**Excellent**
Punctuality/Timely completion of assigned work			Y		
Learning capacity/Knowledge upgradation			Y		
Performance/Quality of work			Y		
Behaviour/Discipline/Team work				Y	
Sincerity/Hard work			Y		
Comment on nature of work done/Area/Topic	He is doing assigned work				
OVERALL GRADE (Any one)	**POOR/AVERAGE/GOOD/VERY GOOD/EXCELLENT**				
Name of Industry Mentor	Dr. Rohit Agrawal				
Signature of Industry Mentor					

Receiving Date	**11-2-2025**	**Name of Faculty Mentor**	**Dr. Rohit Agrawal**	**Sign**	

MONTHLY PROGRESS REPORT (MRP-II)

Name of student	Gaurav Sharma		Department	CSE	
Industry/Organization	MITS		Date/Duration	15/02/2025-15/03/2025	
Criterion	**Poor**	**Average**	**Good**	**Very Good**	**Excellent**
Punctuality/Timely completion of assigned work		Y			
Learning capacity/Knowledge upgradation	Y				
Performance/Quality of work		Y			
Behaviour/Discipline/Team work		Y			
Sincerity/Hard work	Y				
Comment on nature of work done/Area/Topic	He is not regularly updating.				
OVERALL GRADE (Any one)	**POOR/AVERAGE/GOOD/VERY GOOD/EXCELLENT**				
Name of Industry Mentor	Dr. Rohit Agrawal				
Signature of Industry Mentor					

Receiving Date	17-03-2025	**Name of Faculty Mentor**	**Dr. Rohit Agrawal**	**Sign**	

MONTHLY PROGRESS REPORT (MPR) FROM INDUSTRY MENTOR

Name of student	Gaurav Sharma		Department	CSE	
Industry/Organization	MITS Gwalior		Date/Duration	15/03/25 – 15/04/25	
Criterion	Poor	Average	Good	Very Good	Excellent
Punctuality/Timely completion of assigned work			✓		
Learning capacity/Knowledge up-gradation				✓	
Performance/Quality of work				✓	
Behavior/Discipline/Teamwork				✓	
Sincerity/Hard work			✓		
Comment on nature of work done/Area/Topic	He is writing a research paper.				
OVERALL GRADE (Anyone)	POOR/ AVERAGE/ GOOD/ VERY GOOD/ EXCELLENT				
Name of Industry Mentor	Dr. Rohit Agrawal				
Signature of Industry Mentor					

Receiving Date	16-04-2025	Name of Faculty Mentor	Dr. Rohit Agarwal	Sign	