Journal of Cognitive Human-Computer Interaction

♦ ISSN: 2771-1463 Vol: 10, No. 01, 2025 ♦ PP. 16–19

Article DOI: https://doi.org/10.54216/JCHCI.100104

FULL LENGTH ARTICLE

Real-Time Student Identification and Data Retrieval System Powered by Haarcascade and OpenCV

S. Hemamalini[1,*] **J. Beryl Sharon**[1] **M. Dharshini**[1] **M. Indu**[1]

[1] Department of Artificial Intelligence and Data Science, Panimalar Engineering College, India

Emails: hemamalini.phd2020@gmail.com · sharon10bery@gmail.com · dharsh051009@gmail.com · indupriya1920gmail.com

Received: February 05, 2025 Revised: May 05, 2025 Accepted: July 20, 2025 ⋆ Corresponding author

ABSTRACT

Face recognition technology is increasingly integrated into daily life, from unlocking smartphones to taking attendance in classrooms, despite challenges such as lighting, occlusion, and posture variation in real-world scenarios. Therefore, this study aims to develop an automated face recognition system for data retrieval and management using OpenCV. Using a camera, the system records users' photos in real time. Computer vision techniques are then applied, particularly the face identification and recognition functions of the Local Binary Pattern Histogram (LBPH) and the Haar Cascade algorithm, which are implemented using OpenCV. The system correctly recognizes people and makes it easier to handle student information by comparing the faces it detects with a stored database of student photographs. Improved face recognition accuracy, real-time data retrieval, and efficient data management procedures are the main goals. Although the system performed satisfactorily in normal lighting, difficulties with low light were shown to affect detection and recognition accuracy. The primary causes of these constraints were changes in camera quality and lighting. Subsequent developments will concentrate on optimizing the accuracy and overall performance of the system, possibly by incorporating better cameras and more sophisticated processing. The study highlights how computer vision and facial recognition technology can improve data management procedures in a variety of applications. In conclusion, the suggested system effectively makes use of cutting-edge methods for dependable and effective data retrieval.

Keywords: Haar Cascade ▪ OpenCV ▪ Automated Face Recognition ▪ Data Retrieval ▪ Local Binary Pattern Histogram (LBPH) ▪ Computer Vision

1. INTRODUCTION

Face recognition is an application of visual pattern recognition. While face recognition and identification come naturally to humans, a machine must employ sophisticated face recognition algorithms and techniques to carry out this task [1]. It is a significant and increasingly common authentication method. Facial recognition refers to a class of biometric data that uses patterns and biometric information to identify a person based on the head, neck, and face [2]. For the past 20 years, face recognition has been one of the most challenging fields. To create a face recognition system, two approaches are combined: face detection and face recognition. In addition, face recognition technology is rapid, user-friendly, and safe [3].

The purpose of this system is to build a face-recognition-based data retrieval and management system. In this case, a person's face is used to identify them and extract their information from a database [4]. Face recognition technology is becoming more common and is used extensively in many industries. This paper proposes a system that recognizes students' faces from a live webcam feed and displays pertinent

data, such as name, age, and department, if the recognized face is found in the database. The system examines photographs to recognize faces using the Local Binary Patterns Histogram (LBPH) algorithm and recognizes faces in real time using OpenCV's Haar Cascade Classifier. Compared with conventional identifying techniques, the approach provides a more effective, safe, and intuitive solution.

2. RELATED WORKS

Recent studies highlight advancements and challenges in facial recognition technology, especially its applications in security and attendance management systems [5]. A student surveillance system demonstrated reliable performance in natural lighting but struggled in low-light conditions because of camera quality and processing power [6]. Future improvements could involve better hardware integration. A facial recognition system using artificial neural networks achieved high accuracy during evaluations, showing potential for broader applications, including healthcare-oriented detection tasks [7]. Additionally, a robust Python-based system exhibited effective face recognition capabilities, with increased training images improving accuracy [8].

However, limitations such as the use of Principal Component Analysis (PCA) have restricted accuracy to below 90% in some systems [9]. Ongoing research focuses on overcoming challenges such as expression variability and occlusion, with advancements in algorithms including Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM) [2]. The implementation of a web-based attendance management system has shown promise in enhancing tracking accuracy and reducing administrative workload [10]. Future directions include multi-face recognition and offline QR scanning support, further streamlining processes in educational environments [11, 12, 13, 14].

Existing facial recognition systems have notable drawbacks that affect their effectiveness. A significant issue is their performance in low-light conditions, which often results in reduced detection and recognition accuracy because of limitations in camera quality and processing power. Systems relying on PCA may also struggle to exceed certain accuracy thresholds, especially when facial expressions, occlusions, or variations in pose are present. Moreover, many attendance-oriented systems remain dependent on stable network connectivity or controlled environments, limiting their flexibility in real-world educational settings.

3. PROPOSED METHOD

Numerous advancements have been made in face recognition technologies, yet their application in the educational domain remains underutilized. This project addresses this gap by introducing a face-recognition-based student identification system specifically designed to retrieve student data in real time. The system integrates a live camera feed, Haar Cascade face detection, LBPH recognition, and SQLite-based data storage.

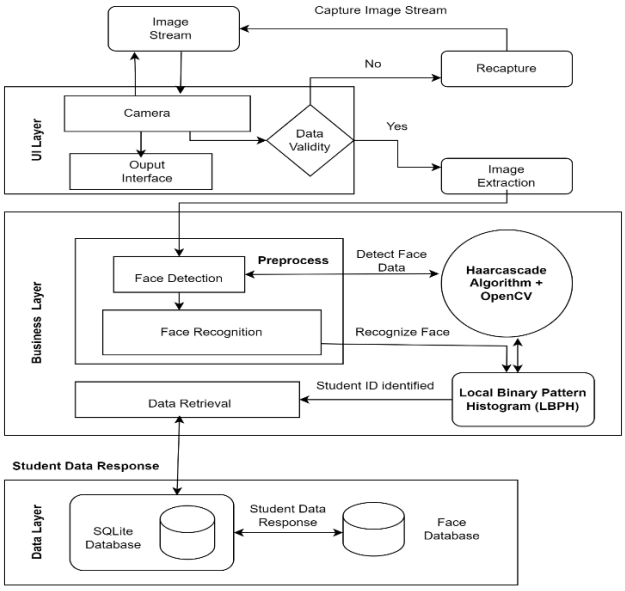

Figure 1. System architecture for the face-recognition-based student identification system.

Figure 1 presents the architecture diagram for the face-recognition-based student identification system and provides a visual overview of how different components interact across three key layers: UI layer, business logic layer, and database layer. In the UI layer, the camera acts as the image input device, capturing live image streams and passing them to the processing modules. The output interface displays the recognized student profile and retrieval results.

In the business logic layer, face detection and face recognition modules handle the core computer vision operations. The Haar Cascade algorithm detects faces from the input frames, while LBPH compares detected faces with the trained face database. The data retrieval component then links the recognized face to the corresponding student record. In the data layer, the SQLite database stores student details, while the face database stores trained facial samples for matching and recognition.

3.1 Dataset Creation

Using a camera to take pictures of students, a dataset of those photographs must be created first. Every student registers by providing their department, age, name, and ID. A webcam takes numerous pictures of the student from different angles and motions, and these details are recorded in the database. The dataset creation process is essential because the recognition model relies on representative images of each registered student.

The captured images are converted to grayscale and labelled with the student's unique ID. These images are stored in a structured dataset directory and are later used for model training. Capturing multiple face images under different facial positions improves recognition robustness and helps the system handle minor variations in pose and expression.

3.2 Model Training

After dataset creation, the LBPH technique is used to train the system to detect and recognize faces. Preparing the training data, training the model, and storing the trained model for

later use are the major steps in the process. By preparing the training data, each image is associated with the correct student ID, allowing the model to learn discriminative face patterns.

LBPH extracts local texture patterns from face images and converts them into histograms that represent facial features. The trained model stores these patterns and compares them with features extracted from live video frames during recognition. This makes LBPH suitable for real-time applications because it is computationally efficient and performs effectively under normal lighting conditions.

3.3 User Detection and Data Generation

This stage involves real-time face detection and recognition. When the system is running, it uses the camera feed to capture live video frames and then detects and recognizes faces. The Haar Cascade Classifier from OpenCV is used for face detection. Haar features are applied to locate face regions quickly, and detected faces are then passed to the LBPH recognizer.

Once a face is recognized, the system maps the predicted ID to a student record. The corresponding student details, including name, age, department, and ID, are retrieved from the SQLite database. This process enables real-time student identification and provides immediate access to stored information.

3.4 Data Management

After detecting and recognizing a face, the student's profile is displayed in real time on the video frame. The system could be extended to incorporate additional functionalities such as generating attendance records, retrieving academic details, or logging the time and date of student access. SQLite is used because it is an embedded SQL database engine that stores an entire database in a single disk file and does not require a separate server process.

The database file format is cross-platform, meaning that databases can be copied between different architectures and systems. Because of these characteristics, SQLite is a practical application file format for lightweight student information systems. It is therefore suitable for storing student records and linking them with recognized face IDs.

4. RESULTS AND DISCUSSION

The model provides facial recognition for student identification and data retrieval in real time. The system begins by collecting user details such as name, age, and department, which are stored in an SQLite database. It then captures grayscale images of the user through the webcam and uses these images for model training. During operation, the Haar Cascade detector locates faces in live frames, while the LBPH recognizer identifies the detected face and retrieves the corresponding student information.

4.1 Model Evaluation

The model's performance is evaluated by calculating accuracy, precision, recall, and F1-score. These metrics provide insights into the model's ability to correctly identify and retrieve student data based on face recognition. The evaluation terms are defined as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives.

Table 1. Results achieved for listed performance measures.

Metric	Description	Results Obtained
Accuracy	Percentage of total correct predictions over all cases	75%
Precision	Proportion of positive identifications that were actually correct	88.1%
Recall (Sensitivity)	Proportion of actual positives that were correctly identified	75%
F1-score	Harmonic mean of precision and recall	88.1%

Table 1 summarizes the achieved performance. The reported accuracy of 75% indicates that the system correctly identifies students in most normal-lighting test cases. Precision of 88.1% shows that most positive identifications correspond to the correct student record. Recall of 75% reflects the proportion of actual positive cases retrieved successfully, while the F1-score summarizes the balance between precision and recall.

4.2 Expected Trends

The algorithm must be trained initially. A dataset containing face photographs of the individuals to be identified is needed. Each photograph should be accompanied by a unique ID, which could be the person's name or ID number. After that, the algorithm uses this data to identify an input image and provide the result. A specific person's image must match the corresponding ID.

In this stage, the facial characteristics are highlighted in an intermediate image produced by applying LBP computation to the original image. The sliding-window concept uses parameters such as radius and neighbors. As more training images are added, the recognizer is expected to improve because it has more representative samples of each student. However, lighting variation, camera resolution, and occlusion can still reduce recognition quality. The system is therefore expected to perform best under normal lighting with clear frontal face

images.

5. CONCLUSION

In conclusion, the automated face recognition system for data retrieval and management effectively uses cutting-edge computer vision techniques to improve the accuracy and efficiency of student identification and attendance management. The system reduces dependency on manual approaches by employing LBPH for face recognition and Haar Cascade for face detection, thereby automating the process. This project can eventually develop into a full-fledged application with improved algorithms designed to meet particular needs for data retrieval.

Future goals include increasing face recognition accuracy, incorporating real-time analytics for effective data management, and creating an intuitive user interface for quick access to student data. Multi-modal biometric authentication could also be added to improve security, and cloud technologies could be used for remote access and scalability. With these improvements, the system can become a strong automated data retrieval solution, greatly increasing the effectiveness of managing student information in a variety of scenarios.

REFERENCES

[1] B. M. Susanto *et al.*, "Face recognition using haar cascade classifier and facenet (a case study: Student attendance system)," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 13, no. 2, pp. 272–284, 2024.

[2] M. Jha *et al.*, "Face recognition: Recent advancements and research challenges," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2022.

[3] S. Bansal, K. Gupta, and A. Sharma, "Enhancing user experience in chatbots: A study on natural language processing techniques," *International Journal of Human-Computer Interaction*, vol. 39, no. 4, pp. 356–367, 2023.

[4] S. Natarajan and D. H. Jayamangal, "Smart attendance monitoring system using facial recognition," *International Journal of Advanced Research in Science, Communication and Technology*, pp. 45–49, 2024.

[5] N. Samriddhi, J. Jaiswal, and P. Dubey, "Face recognition system using haar algorithm," in *2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 2023.

[6] G. Singh and A. K. Goel, "Face detection and recognition system using digital image processing," in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, 2020.

[7] O. A. Al-Shareef and N. M. Gaboua, "Face recognition using deep learning," in *2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, 2023.

[8] N. T. Singh *et al.*, "Student surveillance system using face recognition," in *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, 2023, pp. 843–847.

[9] E. S. K. Siew *et al.*, "Streamlining attendance management in education: A web-based system combining facial recognition and qr code technology," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 33, no. 2, pp. 198–208, 2023.

[10] R. Sharma, V. K. Sharma, and A. Singh, "A review paper on facial recognition techniques," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2021.

[11] R. Kumar, N. Sharma, and P. Verma, "Designing intelligent conversational agents: A review of recent advances," *Journal of Interactive Learning Research*, vol. 33, no. 2, pp. 145–160, 2022.

[12] P. J. Thilaga *et al.*, "Modern face recognition with deep learning," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018.

[13] A. Trivedi *et al.*, "Face recognition based automated attendance management system," *International Journal of Scientific Research in Science and Technology*, pp. 261–268, 2022.

[14] C. Vishwanatha *et al.*, "Face recognition and identification using deep learning," in *2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, 2023, pp. 1–5.