Automated Classroom Attendance System using Machine Learning and Computer Vision

Prayag Thaker, Aarav Desai, Sneh Naik, Diya Contractor, Pradip Kumar Barik
Department of ICT, Pandit Deendayal Energy University, Gandhinagar, India
Email: Prayag.tict22@sot.pdpu.ac.in, Aarav.dict22@sot.pdpu.ac.in, Sneh.nict22@sot.pdpu.ac.in, Diya.cict22@sot.pdpu.ac.in, pradip.barik@sot.pdpu.ac.in

***Abstract*—Marking attendance during classes using traditional methods is a tedious and error-prone process. This traditional approach wastes valuable teaching time and can cause errors which impact student records. To address these challenges, we propose an automated attendance system that applies advanced facial detection and recognition techniques to streamline the process. Our system uses DeepFace for both face detection and recognition, selected after a careful comparison with methods such as Haar cascades and YOLOv9. Among the approaches tested, DeepFace demonstrated the highest reliability and accuracy, effectively identifying students from a single snapshot of the classroom. This automated solution reduces the time and manual effort involved in attendance, making the process more efficient and less intrusive. Our system offers a dependable, user-friendly solution for educators, enhancing routine classroom operations with minimal disruption.**

***Index Terms*—Forest fire detection, fire severity assessment, YOLOv8, Fire Weather Index (FWI), Wildfire management, Artificial Intelligence (AI)**

I. Introduction

Attendance taking at institutions has been an everyday practice, though at times very boring and not relevant. Fig. 1 shows a general picture taken from a classroom for attendance marking [1]. It is important for gauging class participation patterns and meeting institutional requirements. Traditional roll calls, passing of attendance sheets, and swiping of cards by students are very manual, time-consuming, and can disrupt the classroom flow; they also provide room for errors or manipulation. With the advent of artificial intelligence, especially in computer vision and deep learning, there come automated systems that guarantee the process of attendance will be more efficient, save time, and provide reliable tracking of attendance. Among such is face recognition technology, which makes it possible for taking attendance without disrupting the learning environment. After many industrial applications we try to incorporate Machine Learning for attendance automated system.

This paper explores an automated classroom attendance system that uses comprehensive analysis of Haar Cascade, YOLOv9 and DeepFace for face detection and recognition model, a leading-edge technologies recognized for its speed and precision in real-time applications. By taking a single snapshot of the entire classroom, the system can detect and recognize the faces of each student in the frame, marking their attendance instantly. This approach is particularly advantageous in dynamic classroom settings, where capturing

Fig. 1. A general view of classroom attendance

a complete view of the students quickly and accurately is critical to avoid disrupting the lesson. YOLOv9's high-speed processing and advanced face detection capabilities allow it to perform well even under challenging conditions, such as varied lighting, different seating arrangements, and partial occlusions, which are typical in real-life classrooms.

An automated attendance system utilizing Haar Cascade, YOLOv9 and DeepFace are being developed to modernize and streamline attendance tracking in classroom settings. This innovative solution aims to enhance the teaching and learning experience by minimizing administrative tasks and distractions, allowing educators and students to focus on their core objectives.

System Design and Development: The entire flow of the proposed system is divided into a few steps as follows.

- **Dataset preparation:** Curating images to reflect real classroom conditions (varied lighting, seating, facial orientations).
- **Model training:** Fine-tuning Haar Cascade, YOLOv9 and DeepFace for face detection and differentiation in crowded environments.
- **Face recognition algorithm integration:** Creating a database of unique facial features for each student.
- **Balancing speed and accuracy:** Optimizing system performance for real-time attendance tracking. Also au-

tomating the present list of students in excel sheet.
- **System evaluation:** Assessing accuracy and processing time in various classroom setups.

Objectives and Benefits: The main objectives of the attendance system are

- Reliable attendance tracking: Minimizing errors and missed attendance.
- Enhanced classroom experience: Reducing administrative tasks and overhead.
- Increased focus on learning: Allowing educators and students to prioritize teaching and learning.
- Intelligent classroom creation: Leveraging AI-based software for convenient and engaging teaching environments.
- Modernizing attendance processes: Streamlining and automating attendance tasks for efficiency.

The rest of the paper is organized as follows: Section II provides a comprehensive study of existing methods of attendance systems using various primitive methods, Machine Learning, Computer Vision, and their limitations. The methodology used for face detection and recognition for attendance in the classroom is discussed in Section III. Section IV discusses the performance metrics used to evaluate the proposed system. Section V shows the simulation results of the proposed scheme. Finally, conclusions are given in Section VI.

II. Related Work

Work on attendance systems using computer vision (CV) and machine learning (ML) has focused on automating attendance tracking with techniques like LBPH, Haar Cascades, and CNNs, achieving varying levels of accuracy and efficiency, while addressing challenges such as environmental sensitivity, hardware requirements, and data security.
The system proposed by J. Harikrishnan et al. [2] features a robust, user-friendly interface designed for Raspberry Pi, enabling efficient real-time surveillance and attendance tracking with high face detection accuracy using LBPH and Haar cascades. It supports modular deployment in various environments, even in low light, but achieves a maximum recognition accuracy of 74% during real-time use. Another face recognition attendance system by Yanhua Zhan [3] based on real-time video processing offers significant advantages in terms of accuracy, efficiency, and security. However, it also comes with challenges related to cost, computational requirements, environmental sensitivity, privacy concerns, and technical complexity.

The method proposed by Rathod et al. [4] is more economical for large-scale deployment as compared to other biometric systems. This system requires only basic hardware (camera, computer) and a local network. The work by S Rathinamala et al. [5] describes how the system has a robust and user-friendly graphical interface, optimized for Raspberry Pi devices. It requires minimal hardware: a Raspberry Pi and a surveillance camera. High modularity allows for deployment in various environments, including classrooms and labs, with reliable performance even in poor lighting conditions due to effective image pre-processing. This system also automates local as well as cloud data storage, saving time and effort.

Previous research by A Venugopal et al. (2021) uses ensemble of multiple deep learning models (VGG-FACE, Facenet, Openface, DeepFace). Achieves better recognition accuracy through combined model approach. Real-Time Processing: Can process video input in real-time. Capable of identifying multiple students simultaneously. Converts video into frames for efficient processing. Economical: Requires only basic hardware (camera, computer). Simple implementation requirements. Easy to maintain and update database. [6]. But on top of these advantages there were many setback like there were unidentified students left, significant alterations were observed if the angle, distance or lighting changes during the recognition of faces.

Chowdhury et al. [7] propose an idea of an Automatic Class Attendance System using CNN-based Face Recognition, but its performance affected by lighting conditions. Cannot detect faces that are too far from the camera. Needs consistent number of images per person. Requires multiple images (20+ recommended) of each student Training time increases with dataset size. Yet accuracy is only around 91-92% despite increasing training data. Requires high-performance computer for faster training May need periodic dataset updates as well.

CNN systems capture images or videos, perform tasks like feature extraction, mapping, and classification, and learn facial features through training. They can enhance security by integrating with access control systems and CCTV cameras, ensuring authorized access and attendance [8]. OpenCV is a powerful and versatile tool for computer vision tasks, but it requires a good understanding of its capabilities and limitations to be used effectively. [9]. Designed for computational efficiency with focus on real-time applications. It has extensive functionality for computer vision and machine learning tasks. Good integration with other libraries like NumPy for enhanced performance. But TH. Hasan et al. also mentioned that model's performance depends heavily on image quality and computational resources.

The YOLO algorithm is a highly accurate attendance tracking system suitable for enterprises, handling large data sets and integrating with other management systems. It provides statistical data and analysis tools for informed HR decisions. However, it requires significant investment, complex setup, and network connectivity disruptions, and can produce false positives or negatives, affecting performance. [10]. The system offers reliable attendance tracking with 82.44% accuracy in recognition and 100% accuracy in detection. It uses Zoom and Panopto APIs for synchronous and asynchronous online classes, and has an intuitive Python-developed interface. It generates reports for academic performance analysis. However, it relies on network connectivity, may produce false positives, requires regular maintenance, requires training for teachers, and may require further optimization for large datasets. [11]. The Smart Attendance System Based on Face Recognition Techniques offers significant advantages in terms of accuracy, efficiency, and security. However, it also presents challenges related to data security, privacy concerns, and technological limitations. Educational institutions considering this system should carefully weigh these factors and ensure they have the necessary resources and support to implement and maintain it effectively. [12]. Multi-face attendance system by Rath et al. [13], also stated that while the ”Multi-Facial Automated Attendance System using Haar Cascade, LBPH, and OpenCV-Based Face Detection and Recognition” offers significant advantages in terms of accuracy, efficiency, and security, it also presents challenges related to data security, privacy concerns, and technological limitations. Educational institutions considering this system should carefully weight these factors and ensure they have the necessary resources and support to implement and maintain it effectively.

III. Methodology: Face Detection and Recognition Using Haar Cascades, YOLOv9 and DeepFace

We have modeled Haar Cascades, YOLOv9 and DeepFace for face recognition and detection. The numerical comparison of their accuracy has been obtained. Automating classroom attendance using AI-powered face recognition aims to overcome traditional manual method limitations, which are prone to errors, time-consuming, and vulnerable to proxy attendance. Our proposed system leverages YOLOv9 for face detection, Haar Cascades for face localization, and DeepFace for facial recognition, accurately detecting and recognizing student faces in challenging conditions such as large classroom sizes, occlusions, and varying lighting. This integrated solution ensures enhanced accuracy, increased efficiency, and reduced administrative burden, allowing educators to focus on teaching and inspiring students while minimizing errors and discrepancies, promoting fairness and reliability.

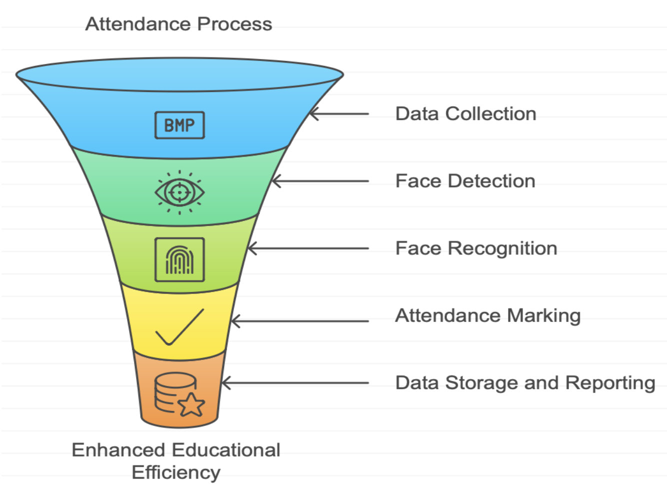

Fig. 2. Overview of detection and recognition process

A. Method-1: Haar Cascades

Haar Cascades utilize Haar features and an integral image representation to efficiently detect faces in images. The key components of Haar Cascades for face detection include:

- **Feature Extraction:** Haar wavelets capture spatial relationships between pixels, which enables the detection of specific facial features.
- **Classifier Training:** The AdaBoost algorithm is employed for training the classifier, selecting the most optimal features for face detection.
- **Cascading:** The classifier undergoes multiple stages, where each stage sequentially rejects non-face regions to improve efficiency.

Although Haar Cascades are less robust compared to deep learning-based detectors, they offer the advantage of fast processing and reliable initial face localization with accuracy greater than or equal to 90% in the following scenarios:

- **Resource-Constrained Environments:** Due to their lightweight nature, Haar Cascades are ideal for devices with limited computational resources.
- **Partial Occlusions:** The method is effective in cases where faces are partially blocked or obscured.
- **Challenging Angles:** Haar Cascades can still detect faces at different orientations, making it suitable for various classroom settings.

B. Method-2: YOLOv9

YOLOv9 (You Only Look Once, version 9) is a state-of-the-art real-time object detection model that excels in multi-object detection tasks. Its architecture offers several advantages for face detection:

- **Convolutional Neural Networks (CNNs):** YOLOv9 uses CNNs to extract features and establish spatial hierarchies for more accurate object localization.
- **Spatial Pyramid Pooling (SPP):** This technique allows YOLOv9 to encode robust features from varied spatial resolutions, improving detection accuracy.
- **Non-Maximum Suppression (NMS):** NMS ensures efficient pruning of redundant bounding boxes, leading to a cleaner and more precise detection.

The combination of these methods makes YOLOv9 capable of detecting multiple faces within a single frame with an accuracy greater than or equal to 95%. This is ideal for:

- **Large Classroom Environments:** YOLOv9 can detect numerous faces in crowded classrooms, making it efficient for attendance tracking in larger groups.
- **Real-Time Attendance Tracking:** Due to its speed, YOLOv9 is suitable for real-time applications, ensuring accurate and quick detection without delays.
- **Low-Latency Applications:** YOLOv9’s optimized architecture makes it well-suited for applications that require immediate feedback, such as real-time attendance marking.

C. Method-3: DeepFace

DeepFace utilizes a deep learning approach to generate facial embeddings. Each detected face is transformed into a 128-dimensional feature vector that encapsulates the unique characteristics of the face. The key components for embedding generation are:

- **Convolutional Neural Networks (CNNs):** DeepFace uses CNNs to extract facial features that are crucial for distinguishing individual faces.
- **FaceNet Architecture:** The FaceNet model is used to optimize the process of embedding generation, ensuring that the embeddings are representative and discriminative.
- **L2 Normalization:** After the embeddings are generated, L2 normalization is applied to scale the embeddings, which improves the comparability of the vectors across different faces.

Similarity Matching:

Once embeddings are generated, they need to be compared with pre-stored embeddings to identify the correct match. This is done through similarity matching techniques:

- **Cosine Similarity:** Measures the cosine of the angle between two embeddings, giving an indication of their alignment. This method is effective for comparing high-dimensional vectors.
- **Euclidean Distance:** Calculates the direct distance between two embeddings, providing a numerical measure of their closeness.

These similarity scores are then used to determine if the detected face matches a registered student.

Thresholding and Decision Making:

To decide whether a detected face matches a registered student, a threshold value is applied to the similarity score:

- **Thresholding:** A predefined threshold (e.g., 0.7 cosine similarity) is used. If the similarity score exceeds this threshold, it is considered a match.
- **Logging Attendance:** If a match is found, the student is logged as "present" and the attendance record is updated accordingly.

IV. Performance Evaluation Metrics for Object Detection

A. Intersection over Union (IoU)

Intersection over Union (IoU) measures the spatial overlap between the predicted bounding box B_{pred} and the ground truth bounding box B_{gt}. It is defined as:

$$\text{IoU} = \frac{|B_{\text{pred}} \cap B_{\text{gt}}|}{|B_{\text{pred}} \cup B_{\text{gt}}|}$$

where $|\cdot|$ denotes the area of the region. **Explanation:** IoU quantifies how well the predicted bounding box overlaps with the actual ground truth bounding box.

B. Precision (P) and Recall (R)

Precision and recall are fundamental metrics for evaluating object detection models.

1) Precision: Precision measures the accuracy of positive predictions. It is defined as:

$$P = \frac{TP}{TP + FP}$$

where:

- TP: True Positives - Correctly detected instances.
- FP: False Positives - Incorrectly detected instances.

2) Recall: Recall assesses the completeness of the detection. It is defined as:

$$R = \frac{TP}{TP + FN}$$

where:

- FN: False Negatives - Missed detections.

C. Average Precision (AP)

Average Precision (AP) quantifies the trade-off between precision and recall by calculating the area under the Precision-Recall curve:

$$\text{AP} = \int P(R)\, dR$$

For discrete points, it is approximated as:

$$\text{AP} \approx \sum_{i=1}^{n} (R_i - R_{i-1})\, P_i$$

where:

- R_i: Recall at the i-th point.
- P_i: Precision at the i-th point.

D. Mean Average Precision (mAP)

Mean Average Precision (mAP) averages AP across all classes:

$$\text{mAP} = \frac{1}{C} \sum_{i=1}^{C} \text{AP}_i$$

where:

- C: Number of classes.
- AP_i: AP for the i-th class.

E. Additional Metrics

1) False Positive Rate (FPR): The False Positive Rate evaluates the proportion of incorrect positive predictions relative to actual negatives:

$$\text{FPR} = \frac{FP}{FP + TN}$$

where:

- TN: True Negatives - Correct negative predictions.

2) True Positive Rate (TPR): Also known as Recall, TPR measures the ratio of correctly detected positives to all actual positives:

$$\text{TPR} = \frac{TP}{TP + FN}$$

3) F1-Score: The F1-Score balances precision and recall, defined as the harmonic mean of both:

$$F1 = 2 \times \frac{P \times R}{P + R}$$

V. Results and Plots

The table compares the performance of DeepFace, YOLOv9, and Haar Cascade models using various metrics. Precision measures the accuracy of positive predictions, while YOLOv9 and DeepFace are highly accurate. Recall quantifies the number of correct positive predictions made out of all positive examples in the dataset. The F1 Score provides a balanced measure of precision and recall, while Haar Cascade functions poorly. The mean Average Precision (mAP@50) measures average precision at 50% overlap, with YOLOv9 performing well. The R^2 Score indicates how well predictions match actual values, with DeepFace displaying the best fit. The mean squared error (MSE) measures the average squared difference between predicted and actual values. True Positives (TP) and False Positives (FP) are the most applicable metrics. The confusion matrix and precision-recall curve show DeepFace as the best performer, YOLOv9 is balanced, and Haar Cascade performs poorly.

VI. Conclusion:

In conclusion, our research pioneers a cutting-edge approach to attendance tracking, harnessing the power of machine learning and computer vision. By comparing Haar Cascades, YOLOv9, and DeepFace, we've created a robust facial recognition system that revolutionizes the way we take attendance. Our findings reveal that each model has its unique strengths and limitations. Haar Cascades offers speed and efficiency, but compromises on accuracy. YOLOv9 excels in real-time detection, making it ideal for large-scale applications. DeepFace, with its unparalleled accuracy, proves to be the most reliable choice, albeit requiring more computational resources.

Expanding the scope of this attendance system to include large-scale applications, integrating it with school or university infrastructure for automated attendance marking and monitoring could be a possible future direction of this work.

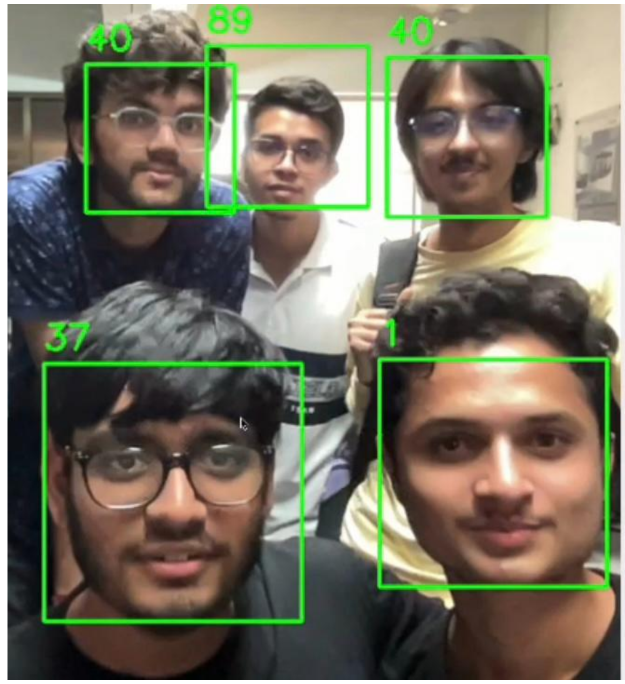

Fig. 3. First Trial Run

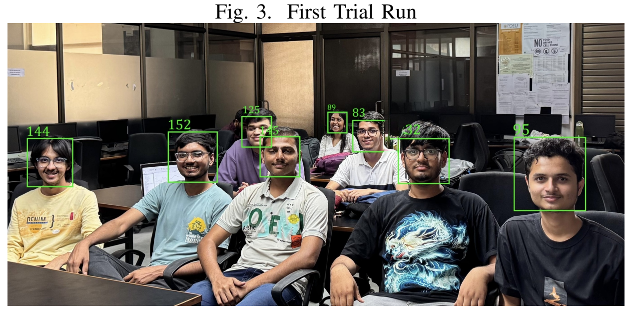

Fig. 4. Classroom Implementation

Name	Date	Time
32	2024-11-07	17:39:16
83	2024-11-07	17:39:17
89	2024-11-07	17:39:18
95	2024-11-07	17:39:19
125	2024-11-07	17:39:20
144	2024-11-07	17:39:21
152	2024-11-07	17:39:22
145	2024-11-07	17:39:23
88	A	
90	A	
100	A	
101	A	
114	A	
120	A	
130	A	

Fig. 5. Attendance Sheet Updation

TABLE I
PERFORMANCE EVALUATION OF HAAR CASCADE, YOLOv9, AND DEEPFACE MODELS BASED ON KEY METRICS

Metric	Haar Cascade	YOLOv9	DeepFace
Precision	1.0	0.9641	1.0
Recall	0.0345	0.7871	1.0
F1 Score	0.0667	0.866	1.0
mAP@50	N/A	0.8894	N/A
R^2 Score	0.0	N/A	0.0
MSE	0.875	N/A	0.0234
True Positives (TP)	10	167	60
True Negatives (TN)	0	N/A	N/A
False Positives (FP)	0	30	0
False Negatives (FN)	280	38	0
Confusion Matrix	$[0, 0], [280, 10]$	$[167, 30], [38, 0]$	$[60]$
Precision-Recall Curve	Low recall, high precision	High recall, moderate precision	Perfect precision and recall

REFERENCES

[1] N. Available, "Classroom and students." Stock photo, 2022. Accessed: 2022.

[2] J. Harikrishnan, A. Sudarsan, A. Sadashiv, and R. A. Ajai, "Vision-face recognition attendance monitoring system for surveillance using deep learning technology and computer vision," in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, pp. 1–5, 2019.

[3] H. Yang and X. Han, "Retracted: Face recognition attendance system based on real-time video processing," *IEEE Access*, vol. 8, pp. 159143–159150, 2020.

[4] H. Rathod, Y. Ware, S. Sane, S. Raulo, V. Pakhare, and I. A. Rizvi, "Automated attendance system using machine learning approach," in *2017 International Conference on Nascent Technologies in Engineering (ICNTE)*, pp. 1–5, 2017.

[5] R. S, K. A, A. R, and C. K. Kannan, "Automatic attendance monitoring system using lbph and haar algorithm," in *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, pp. 1604–1608, 2023.

[6] V. A, R. R. Krishna, and R. V. U, "Facial recognition system for automatic attendance tracking using an ensemble of deep-learning techniques," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1–6, 2021.

[7] S. Chowdhury, S. Nath, A. Dey, and A. Das, "Development of an automatic class attendance system using cnn-based face recognition," in *2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*, pp. 1–5, 2020.

[8] K. V. Kumars, "Non-Intrusive facial recording system with convolution neural network (CNN) for attendance assimilation."

[9] R. T. Hasan and A. B. Sallow, "Face detection and recognition using opencv," *Journal of Soft Computing and Data Mining*, vol. 2, p. 86–97, Oct. 2021.

[10] Y. Zhang, "Design of a remote meeting real-time attendance system based on the yolo algorithm," in *2021 IEEE International Conference on Industrial Application of Artificial Intelligence (IAAI)*, pp. 330–336, 2021.

[11] J. R. M. Lopez, J. V. H. Estabillo, C. C. Paglinawan, and E. D. Dimaunahan, "Utilizing yolov5 and python for integrated attendance checking in hyflex classes in mapua university," in *2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE)*, pp. 838–842, 2023.

[12] A. Al-sabaeei, H. Al-khateeb, A. Al-basser, H. Al-Sameai, M. Al-shameri, and M. Derhem, "Smart attendance system based on face recognition techniques," in *2021 1st International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, pp. 1–6, 2021.

[13] S. Rathi, S. Kapadnis, G. Sonar, G. Sonar, and S. R. Vispute, "Multi-facial automated attendance system using haar cascade, lbph, and opencv-based face detection and recognition," in *2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, pp. 1–4, 2023.