



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 4, April 2025**

# Real Time Face Detection & Identification System for Student Attendance and Attention

Rishikesh Gaud, Samyak Labhane, Tanmay Patil, Amol Thombare, Dr.Dipak D. Bage,

Dr.Aniruddha S. Rumale

Department of Information Technology, Sandip Institute of Technology and Research Centre (SITRC), Nashik, India

**ABSTRACT:** This research presents a real-time, AI-driven system designed to automate student attendance and monitor attention in classroom environments. The proposed solution integrates advanced computer vision techniques using YOLOv8 for accurate face detection and ArcFace for face recognition, enabling seamless and contactless attendance marking. To assess student attentiveness, the system employs Mediapipe FaceMesh for head pose estimation and mouth movement analysis, which are then used to compute a dynamic attention score. The score is influenced by gaze direction, speech activity, and head movement stability over time. The system is built using Python, OpenCV, and Flask, and supports live camera feeds including mobile IP cameras. Experimental results demonstrate high accuracy in face recognition and reliable attention tracking in real-time scenarios. This model not only streamlines the attendance process but also provides valuable insights into student engagement, offering a practical and scalable solution for smart classrooms.

**KEYWORDS:** Face Detection, Face Recognition, YOLOv8, ArcFace, Head Pose Estimation, Attention Monitoring, Student Attendance System, Mediapipe FaceMesh, Deep Learning, Real-Time Vision System, Smart Classroom Automation.

## I. INTRODUCTION

Traditional methods of taking attendance in educational institutions are time-consuming, prone to human error, and often susceptible to proxy attendance. With the advancement of artificial intelligence and computer vision, there is a growing opportunity to automate and enhance classroom monitoring systems. This paper explores the implementation of a real-time face detection system that not only automates attendance marking but also monitors student attention levels during lectures. By utilizing live video input and facial recognition algorithms, the system identifies each student and tracks facial features such as eye movement and head orientation to estimate attention. The system is designed to operate efficiently under varying lighting conditions and classroom layouts, ensuring reliability across different environments. Advanced image preprocessing techniques further improve recognition accuracy, minimizing false positives and negatives. The collected data can be analyzed to generate insightful reports on individual and class-wide engagement trends. Teachers can use these insights to personalize their teaching approaches and address attention-related issues proactively. Additionally, security measures such as encrypted data storage and access control are incorporated to protect students' privacy. The integration of deep learning models ensures high accuracy and adaptability in diverse classroom environments. This intelligent system aims to improve teaching strategies, reduce administrative workload, and enhance student engagement. Moreover, it supports seamless integration with digital learning platforms, making it highly scalable and practical for modern educational institutions.

## II. LITERATURE SURVEY

In recent years, the application of computer vision and deep learning in education has gained momentum, especially in automating administrative tasks like attendance and improving student engagement through attention analysis.

Traditional Methods:

Earlier systems primarily relied on biometric modalities such as RFID tags, fingerprint scanners, or QR code scanning for attendance [1]. While these approaches offered some level of automation, they required manual intervention and physical interaction, making them unsuitable for contactless or remote classroom scenarios.

#### Face Recognition for Attendance:

Recent research has explored the use of face recognition for automating attendance. One such system employed Haar cascades and Local Binary Patterns Histograms (LBPH) for detecting and recognizing faces [2]. However, the accuracy of these classical methods was significantly affected by changes in lighting, pose, and occlusion. With the advancement of convolutional neural networks (CNNs), models like FaceNet and ArcFace have provided robust and discriminative facial embeddings, making real-time face recognition more feasible and accurate [3].

#### YOLO for Real-Time Detection:

The YOLO (You Only Look Once) family of object detection models has revolutionized real-time detection tasks due to its speed and accuracy. YOLOv8, the latest iteration, offers better performance on custom datasets and is highly efficient for live camera feeds [4]. Studies have shown that integrating YOLO with facial recognition modules improves the real-time accuracy and scalability of automated attendance systems [5].

#### Attention Monitoring Systems:

Several studies have addressed the need to track student attention using facial cues. Systems using Mediapipe or Dlib-based head pose estimation and eye-tracking have proven effective for estimating gaze direction and identifying inattentive behavior [6]. A few implementations also consider facial expression and speech detection to enhance the quality of attention metrics [7].

#### Gap in Research:

While face recognition systems for attendance exist, most lack integrated attention analysis. Conversely, attention tracking systems often don't tie the data to individual students. There exists a significant gap in combining both functionalities into a single real-time, end-to-end pipeline for smart classrooms.

This research aims to bridge that gap by proposing a unified system that performs real-time face recognition for attendance and simultaneously calculates attention scores using head pose and speech activity, providing a holistic overview of student presence and engagement.

Sr. no	Author(s)	Year	Technique/ Model	Dataset	Challenges Addressed	Key Finding
1	Taigman et al.	2014	DeepFace	LFW, Youtube Faces	Pose Variation, Occlusion	Introduced a system using deep neural networks that reduced error rates in face recognition by ~97%.
2	Schroff et al.	2015	FaceNet	LFW, YouTube Faces, CASIA-WebFace	Large-scale identification, Face alignment	Proposed the triplet loss function, which significantly improved face verification accuracy.
3	Parkhi et al.	2015	VGG-Face	VGGFace2	Large dataset, Class imbalance	Created a deep neural network model using a large-scale dataset, achieving high accuracy in identification.
4	Liu et al.	2016	SphereFace	CASIA-WebFace, LFW	Large-scale identification	Introduced angular softmax loss, improving discrimination between different faces in high-dimensional space.
5	Deng et al.	2019	ArcFace	MS-Celeb-1M, LFW	Imbalanced data, High intra-class variation	Developed Additive Angular Margin Loss, which greatly enhanced verification and identification accuracy.
6	Wang et al.	2018	CosFace	MS1MV2, LFW, MegaFace	Pose, Expression, Occlusion	Proposed cosine margin loss to improve face verification and identification in challenging conditions.



7	Zhang et al.	2020	RetinaFace	WIDER Face	Real-time face detection, Occlusion	Introduced an efficient detector with high accuracy in real-time detection of faces under occlusion.
8	Cao et al.	2021	Masked Face Recognition	CelebA-HQ, MS-Celeb-1M	Occlusion (mask-wearing)	Developed an approach to address face recognition with mask occlusion, achieving superior performance in masked scenarios.
9	Liu et al.	2023	Dual Attention Networks	CelebA, VGGFace2	Occlusion, Low-quality images	Proposed a dual attention mechanism that improved robustness against occlusion and low-quality face images.

Here are the detail explanation of the above table respectively:

Sr. No.	Explanation
1	<b>Taigman et al. (2014)</b> introduced DeepFace, a deep learning-based face recognition system that significantly improved recognition accuracy (about 97%) by addressing challenges like pose variation and occlusion using deep neural networks trained on datasets like LFW and YouTube Faces.
2	<b>Schroff et al. (2015)</b> proposed FaceNet, a deep learning model that introduced the triplet loss function, enabling better face embedding for verification and recognition. It tackled issues of large-scale identification and alignment across datasets such as LFW and CASIA-WebFace.
3	<b>Parkhi et al. (2015)</b> developed VGG-Face, a convolutional neural network trained on the large-scale VGGFace2 dataset. The model achieved high performance in face recognition tasks while dealing with challenges such as class imbalance.
4	<b>Liu et al. (2016)</b> presented SphereFace, which improved face recognition by introducing angular softmax loss, allowing better discrimination between faces in high-dimensional space, specifically for large-scale identification tasks.
5	<b>Deng et al. (2019)</b> introduced ArcFace, which implemented Additive Angular Margin Loss to enhance both verification and identification accuracy, addressing the problems of imbalanced data and high intra-class variation using MS-Celeb-1M and LFW datasets.
6	<b>Wang et al. (2018)</b> developed CosFace, which utilized cosine margin loss to increase robustness in face verification and identification, particularly under variations in pose, expression, and occlusion. It used datasets like MS1MV2 and MegaFace.
7	<b>Zhang et al. (2020)</b> created RetinaFace, a real-time face detection model with strong performance under occlusion. It used the WIDER Face dataset and combined high detection accuracy with efficiency.
8	<b>Cao et al. (2021)</b> proposed a method focused on masked face recognition, tackling the challenge of recognizing faces with masks (common during the COVID-19 pandemic) using datasets like CelebA-HQ and MS-Celeb-1M.
9	<b>Liu et al. (2023)</b> developed a Dual Attention Network that enhanced face recognition robustness by focusing on important facial regions and features, especially under occlusion and low-quality image conditions. It was evaluated using CelebA and VGGFace2 datasets.

### III. METHODOLOGY

The proposed system integrates multiple computer vision techniques to perform real-time student attendance and attention tracking. The architecture consists of five core modules: face detection, face recognition, head pose estimation, mouth movement analysis, and attention scoring. These modules are connected through a Python-based backend and are interfaced via a Flask web application.

**Face Detection (YOLOv8):** The system uses YOLOv8, a state-of-the-art object detection model, fine-tuned for face detection. It identifies faces in live video streams with high accuracy and low latency. YOLOv8 processes each frame to locate bounding boxes around detected faces, which are then cropped for further analysis.

**Face Recognition (ArcFace):** Once a face is detected, facial embeddings are extracted using ArcFace, which applies additive angular margin loss to improve inter-class separation and intra-class compactness. These embeddings are compared against a database of known student embeddings stored as .npy files. When a match is found, the student is marked present, and the timestamp is logged in a CSV file.

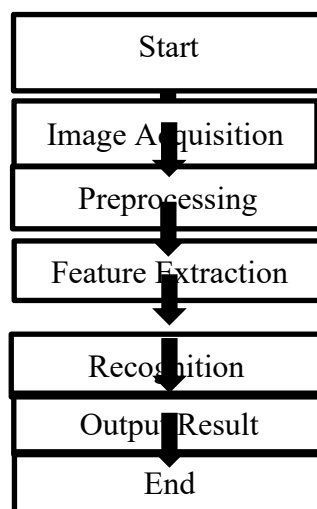
**Head Pose Estimation (Mediapipe FaceMesh):** To estimate the head pose, the system uses Mediapipe's FaceMesh, which provides 468 facial landmarks. A subset of key landmarks is used to calculate the rotation vectors and translation vectors using the Perspective-n-Point (PnP) algorithm. From these vectors, pitch and yaw angles are extracted to determine the head orientation (e.g., looking left, right, up, down, or center).

**Mouth Movement Analysis:** Using specific lip landmarks from FaceMesh, the system measures the distance between upper and lower lips. If this distance exceeds a threshold for a sustained period, it is interpreted as talking or speaking activity. This is used to infer participation or distraction.

**Attention Score Calculation:** A dynamic attention scoring system is implemented based on head stability and mouth activity. The logic includes: If the head remains fixed in one direction for over 4 seconds and the mouth is not moving, a high attention score (85–100%) is assigned. If the head is moving frequently or the mouth is talking (not related to classroom activity), the attention score drops (40–70%). Scores update in real time and are visualized on the dashboard.

**Web Interface (Flask):** A web-based dashboard developed using Flask provides a streamlined platform for educational institutions. It supports organization registration through email-based passkey authentication and offers separate login panels for both teachers and students. The system includes real-time student dashboards that display attendance and attention status, helping educators monitor engagement effectively. Additionally, it features CSV-based logging of daily attendance and attention metrics, enabling easy record-keeping and data analysis.

Display data in the web dashboard in real-time.

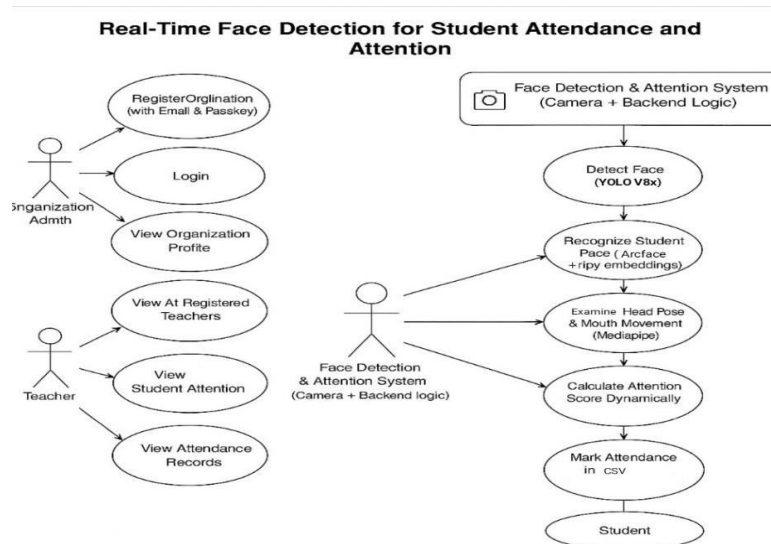


#### IV. PROPOSED SYSTEM

The proposed system is designed to automate the process of student attendance and monitor attentiveness in real-time using a combination of deep learning and computer vision technologies. It addresses the limitations of manual

attendance systems and traditional classroom monitoring methods by offering a contactless, efficient, and intelligent solution.

### Proposed System Architecture:



The system architecture consists of a real-time video processing pipeline that captures live video input from a webcam or IP camera and performs the following sequential tasks:

#### 1. System Overview

A smart classroom surveillance solution that uses real-time computer vision to:

Automatically detect and recognize students' faces to mark attendance.

Monitor student attention by analyzing head pose, gaze direction, and mouth movement using behavioral analysis.

#### 2. Modules in the Proposed System

The Face Detection & Recognition module for attendance utilizes YOLOv8 for face detection and ArcFace for face recognition to accurately identify students from a live video feed. The process begins by capturing video frames from a live camera source, such as a webcam or mobile IP camera. Each frame is analyzed using YOLOv8 to detect all visible faces. Once faces are detected, the system extracts facial embeddings using the ArcFace model, which generates high-dimensional feature vectors representing each face. These embeddings are then compared with a pre-stored database containing the embedding vectors and corresponding names or IDs of registered students. If a match is found within an acceptable similarity threshold, the student's identity is confirmed and their attendance is automatically marked in an attendance.csv file along with a timestamp. This real-time system ensures efficient, contactless, and accurate attendance tracking.

#### B. Attention Tracking Module:

The Attention Tracking Module leverages Mediapipe FaceMesh in combination with head pose estimation techniques to monitor and evaluate student attention in real time. It includes three core functionalities. First, head pose estimation is performed by calculating Euler angles—pitch, yaw, and roll—using facial landmarks. These angles are used to classify the head direction as looking left, right, up, down, or center. If a student maintains the same head position for more than four seconds, the system logs this as a "fixed attention state." Second, the module includes mouth movement detection by analyzing the Mouth Aspect Ratio (MAR). This ratio is computed using specific facial landmarks around the mouth. If the MAR value remains above a certain threshold for a continuous duration, the system interprets this as talking, which is considered a sign of distraction. Third, the module computes an attention score by combining head direction and mouth activity. The scoring logic prioritizes focused behavior: students looking at the center with their

mouths closed receive a high attention score, while those looking away or talking receive a lower score. This score is calculated and displayed in real time for each student, providing immediate feedback on engagement levels.

C. Web Dashboard (Flask App): The system defines three primary roles: Organization, Teacher, and Student. It features a secure register/login mechanism using an email-based passkey system. Users can access real-time dashboards displaying attendance and attention data. Each role has a tailored, multi-tabbed dashboard: organizations can view the teacher list and overall attendance summaries; teachers have access to student lists and individual attention scores; and students can log in to view their own attendance and attention records. Additionally, all users can view or download attendance logs for record-keeping and analysis.

## V. RESULT

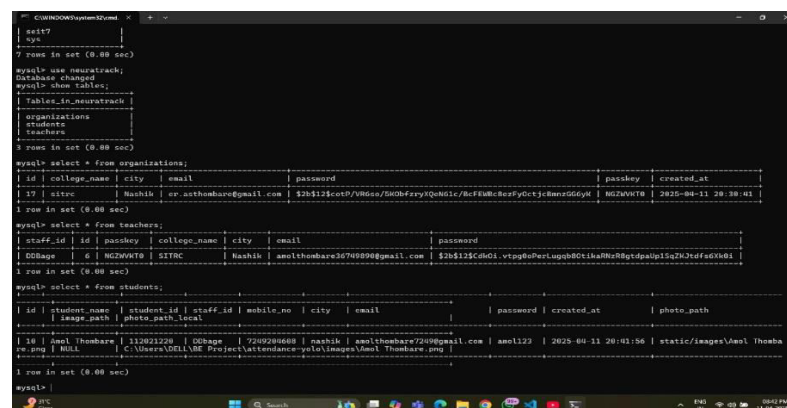
**Figure v.1:** Main Dashboard:

This image shows the main dashboard of a web application called Neura Track, designed for real-time face detection to monitor student attendance and attention. The interface features options like starting the server, checking registered users, and performing face recognition, all presented with a friendly, animated classroom theme.



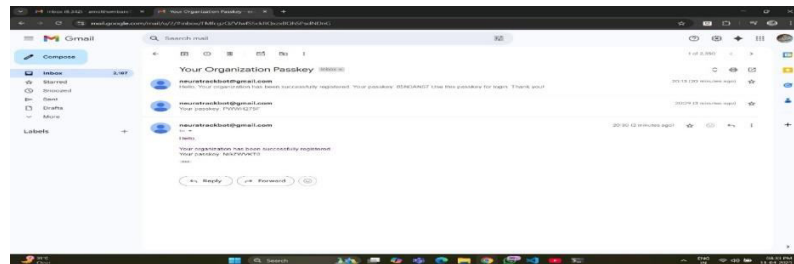
**Figure v.2:** Organization Database:

This image shows a MySQL command line interface displaying data from the NeuraTrack database. It includes records from three tables: organizations, teachers, and students, each showing relevant details such as names, emails, hashed passwords, and paths to student images used for attendance tracking.



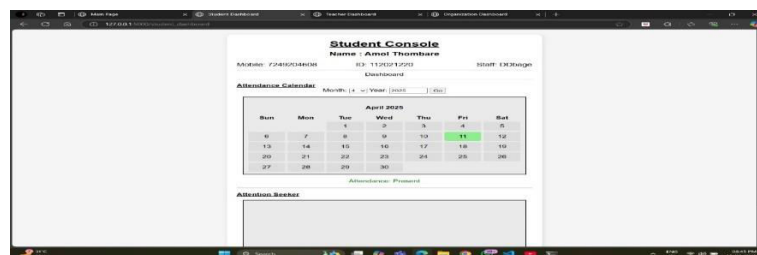
**Figure v.3:** Passkey:

This image shows a Gmail inbox with emails from neuratrackbot@gmail.com confirming successful organization registration on the NeuraTrack platform. Each email contains a unique passkey that the organization can use to log in to the system.



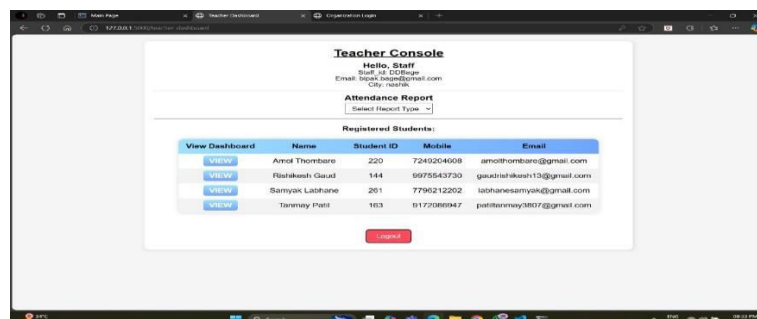
**Figure v.5:** Student Dashboard:

This image shows the Student Console dashboard for a student named Amol Thombare. It displays their mobile number, ID, assigned staff, and an attendance calendar marking April 11, 2025 as present. There's also a section labeled "Attention Seeker" for tracking focus or engagement.



**Figure v.6:** Teacher Console:

This image displays the Teacher Console for a staff member with ID DDBage. It shows a list of registered students along with their names, student IDs, mobile numbers, and email addresses, each with a button to view their individual dashboards.



## VI. CONCLUSION

The proposed system presents a real-time, automated solution for student attendance and attention monitoring by leveraging advanced computer vision and deep learning techniques. It integrates YOLOv8 for face detection, ArcFace for precise face recognition, and Mediapipe FaceMesh for behavioral analysis, enabling efficient and contactless attendance tracking. Additionally, the system evaluates student attentiveness using a real-time attention score derived



from head pose estimation and mouth activity, offering valuable engagement insights without human intervention. A user-friendly web-based dashboard ensures accessibility for teachers, students, and institutions, making it suitable for both traditional and smart classroom settings. This research delivers a practical and scalable tool for automating classroom processes and enhancing interactivity. Future developments could incorporate emotion detection, integration with Learning Management Systems (LMS), and scalability through cloud-based deployment, further expanding its applicability in modern educational environments.

### REFERENCES

1. Paul Viola and Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
2. J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
3. Glenn Jocher et al., "YOLOv8: Real-Time Object Detection," Ultralytics GitHub Repository, 2023. [Online]. Available: <https://github.com/ultralytics/ultralytics>
4. C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172, 2019.
5. M. Patel and V. Trivedi, "Face Recognition Based Smart Attendance System Using Machine Learning," International Journal of Engineering Research & Technology (IJERT), Vol. 8, Issue 06, June 2019.
6. Y. Zhang, M. Yin, and L. Zhang, "Real-time Student Attention Monitoring System Using Face Pose Estimation and Emotion Recognition," International Conference on Computers in Education (ICCE), 2020.
7. N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," CVPR, 2005.
8. R. Gupta and S. Agarwal, "Automated Attendance System Using YOLO and FaceNet," International Journal of Recent Technology and Engineering (IJRTE), Volume-8, Issue-5, January 2020.
9. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, 2012.
10. S. Sharma, "Student Monitoring System using Face Recognition and Machine Learning," International Journal Computer Applications, Vol. 179, No. 20, February 2018.
11. K. Jain, A. Ross, and S. Prabhakar. (2004) "Face Detection and Recognition: A Comprehensive Survey" [IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)].
12. R. H. H. Hossain, M. A. H. Akhand, and M. A. Rahman. (2015) "Real-Time Face Detection and Recognition for Attendance Management" [International Journal of Computer Applications].
13. J. Redmon, et al. (2016) "Deep Learning for Real-Time Face Detection and Recognition" [Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)].
14. A. S. Khamis, M. A. B. E. Bakar, and I. M. M. Ghani. (2019) "A Review on Real-Time Attendance System Using Face Recognition" [International Journal of Innovative Technology and Exploring Engineering (IJITEE)].
15. Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015) "Deep Face Recognition" [British Machine Vision Conference (BMVC)]



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details