Realtime Monitoring Using Face Recognition

Ananth Karthick.M
Artificial intelligence and Data science
Panimalar Institute of Technology,
Chennai, India.
ananthkarthick07@gmail.com

Karthikeyan S R
Artificial intelligence and Data science
Panimalar Institute of Technology,
Chennai, India.
srkarthik27@gmail.com

Yuvaraj A
Artificial intelligence and Data science
Panimalar Institute of Technology,
Chennai, India
yuvi.pit.ai@gmail.com.

Mrs. C M Hilda Jerlin
Assistant Professor
Artificial intelligence and Data science
Panimalar Engineering College,
Chennai, India.
ruthjerry97@gmail.com

Dr. .C. Gnanaprakasam,
Associate Professor
Artificial intelligence and Data science
Panimalar Institute of Technology,
Chennai, India.
cgn.ds2021@gmail.com

Dr. T. Kalai Chelvi,
Professor and Head
Artificial Intelligence and Data Science
Panimalar Institute of Technology,
Chennai, India.
ai_dshod@pit.ac.in

***Abstract*-- This project presents a real-time student monitoring system using deep learning for accurate and efficient face recognition in educational environments. It integrates MTCNN for face detection, FaceNet for feature extraction, and SVM for lightweight classification. The system tracks student movements across campus locations like classrooms, labs, and restrooms, providing real-time updates to administrators and enhancing security and accountability. Designed for low-resource devices, it operates efficiently under varied lighting and environmental conditions. Future enhancements include scaling for larger populations, improving real-time responsiveness, and deploying on edge devices for cost-effective scalability. This solution promotes smarter, safer, and more organized educational spaces through AI-driven monitoring.**

***Keywords*: Face Recognition, Deep Learning, MTCNN, FaceNet, Machine Learning, Real-Time Detection, SVM.**

I. Introduction

In modern educational environments, tracking student movements is essential for ensuring safety, accountability, and operational efficiency. Traditional methods like roll calls and manual registers are inefficient, error-prone, and easily manipulated. With advancements in artificial intelligence and facial recognition, automated, non-intrusive monitoring systems now offer real-time insights into student whereabouts.

Unlike conventional systems focused solely on attendance, this project introduces a solution for continuous location monitoring. Using MTCNN for accurate face detection, FaceNet for robust feature extraction, and SVM for classification, the system automatically identifies students through live camera feeds placed at strategic campus locations. Each movement is logged with identity, location, and timestamp in a structured SQL database, creating a comprehensive timeline of student activities and enabling detection of behavioral anomalies like delays or avoidance patterns.

This dynamic monitoring eliminates loopholes such as proxy attendance, improves transparency, and reduces administrative workload. The system is designed for scalability, suitable for both small classrooms and large campuses, and operates efficiently even on devices with limited computational resources.

Future enhancements include predictive analytics to assess absenteeism risks and student disengagement, improving both academic integrity and well-being. Ethical considerations such as privacy, informed consent, and compliance with data protection regulations like GDPR are integral to the system's design.

In summary, this AI-driven monitoring system moves beyond static attendance, providing a scalable, efficient, and secure solution for modern educational institutions. By integrating advanced face recognition technologies and real-time data analytics, it sets the foundation for smarter, safer, and more organized learning environments.

II. Literature Review

Facial recognition technology is now being used in a lot of everyday situations—whether it's for surveillance, marking attendance in schools, or ensuring security in workplaces. Over the years, many researchers have tried to make these systems more accurate and faster, especially when they're used in places like classrooms where conditions can be unpredictable. Our project, called Face Spotter Vision Hub, is shaped by many of these earlier works. We focused mainly on studies that deal with face detection, deep learning techniques, real-time processing, and privacy concerns, since those were the most relevant to what we wanted to build.In this section, we'll go through some of the important works we reviewed during development.

To start with, Alruwais and Zakariah [1] developed a system that can recognize students and observe their activity in e-learning platforms using deep learning. Their system worked quite well for virtual classes, tracking faces and engagement. But one limitation was that it didn't handle the offline situations where things like background noise, different lighting, or students moving around are common. Our system focuses more on real-time recognition in physical settings, which made us think differently about how to solve those issues.

Zhang et al. [2] introduced a multitask cascaded convolutional network (MTCNN) that can do both face detection and face alignment at the same time. This model became very useful for many developers because it worked well even when faces were turned at different angles or not fully visible. We found this useful since, in classrooms or public spaces, people rarely sit still and face the camera directly. This inspired the design of our face detection module.

A major shift in facial recognition came with FaceNet by Schroff et al. [3], which created a way to turn face images into embeddings that can be compared using distance rather than classification. This means the system can figure out how similar two faces are, even if they're not seen before. We liked this idea because it avoids training for every single person in advance and works better in open-world environments like schools.

On the older side, the concept of support vector machines (SVMs) by Cortes and Vapnik [4] was also important in earlier face recognition models. Though deep learning now dominates this field, SVMs still provide a solid base for understanding how classification models separate data. Some face recognition systems still use SVMs after extracting features with CNNs. We studied this mainly to understand how hybrid systems work.

For monitoring unusual patterns over time, Malhotra et al. [5] proposed an LSTM-based model for time series anomaly detection. While this wasn't exactly about face recognition, it gave us ideas about tracking behavioral patterns — like spotting students who are constantly absent or inactive. If we expand our project later, time-based tracking like this can be very useful.

Transformers have been used for a similar purpose too. Zhao et al. [6] used transformer-based behavior models in educational systems to track students continuously. Transformers pay attention to important patterns across time and were able to detect unusual behavior quite accurately. Although our current system is event-driven, their method gave us some ideas for adding temporal analysis later.

Privacy is a big concern when dealing with facial data. Kim, Choi, and Kwak [7] developed a federated learning approach where the model is trained across different devices without sharing raw images. This is useful in sensitive environments like schools. We didn't use federated learning directly, but it encouraged us to handle face data more carefully and consider privacy at every step.

Another useful work was by the Sun et al. [8], who introduced a self-supervised method using facial landmarks. It helped train recognition models without manually labeling data, which saves a lot of time and effort. In practical use, systems like ours often don't have perfect datasets, so learning from unlabeled data is a good alternative.

In terms of real-time systems, Sharma and Patel [9] implemented a YOLOv7-based attendance system that could detect and recognize faces quickly. Their results showed that fast and accurate face detection is achievable, even on regular hardware. We were especially inspired by their success in classrooms, which made us confident to use YOLO-based methods for our project.

Chen and Liu [10] added something new to this area by tracking students across multiple cameras using graph neural networks. That allowed the system to follow a student from one room to another. We're not using multi-camera systems right now, but their work opened up ideas for how large institutions could implement better tracking.

Ahmed and Islam [11] built a face recognition attendance system using simple tools like a Raspberry Pi and a CNN. Their goal was to create a low-cost solution, which we also found meaningful because many schools don't have the budget for high-end AI systems. It confirmed that face-based systems can be built with basic resources.

Balakrishnan et al. [12] added alert systems to facial recognition, which notify users when unknown faces are seen. That extra feature made their system more active than just a recognizer. We are planning to include similar alert options to help users respond to problems immediately.

Some researchers also tried combining face recognition with other technologies. For example, Park et al. [13] combined BLE beacons with facial recognition to improve tracking. Though we didn't use BLE due to hardware limitations, their approach showed that combining multiple data types could help reduce false results.

As for the tools used to build such systems, TensorFlow [19], Keras [20], and OpenCV [21] were common across almost all recent projects. These tools made it easier to handle deep learning models, video feeds, and face detection in real-time. We also relied on them during our development phase.

From an ethical point of view, Cavoukian [16] proposed the "Privacy by Design" principle which says privacy should be considered from the start of any system. We thought it was important to take user privacy seriously, so we made sure to encrypt facial data and avoid storing anything longer than necessary. The General Data Protection Regulation (GDPR) [22] also outlines strict rules about handling personal information, and we used it as a guide to help shape how our system manages data securely.

In terms of ethics, Jones and Silver [17] pointed out that constant monitoring in schools can affect how students behave and may even invade their privacy. That really stood out to us, especially since our system is designed for classroom use. Gonzalez and Park [18] also emphasized that ethical AI shouldn't just be about having guidelines—it needs to be built directly into the software itself. Taking that into account, we added opt-out features and made sure that only authorized admins could access or manage sensitive data.

III. Existing System

So, in most colleges or offices, attendance is still done the old-school way. Like, they either call out names or pass around a register to sign. Honestly, it takes time and sometimes, people just mark for others. That's a big problem, especially in big classes. A few places now use ID cards or fingerprint scanners. But yeah, touching a fingerprint scanner all the time? Not great after COVID. Some newer systems try to use face recognition, which sounds cool, but most of them don't really work well in normal conditions. Like, if the lighting's bad or someone's not facing the camera exactly right, they mess up. Also, many of those systems don't protect your face data. It just gets stored, often not even encrypted. That's kind of scary if you think about it. We figured there's still a big gap—there needs to be a better, safer, and cheaper way to track people accurately, without them needing to touch anything or worry about where their data ends up.

IV. Proposed System

We were looking for a smarter way to take attendance because calling out names or using ID cards just takes too much time and feels pretty old-school. That's why this system was built to use a regular camera to quickly catch faces and figure out who's there without any hassle. It works live, without needing anyone to touch anything, which is way more hygienic—something people care about more now after COVID. The goal was to build something that's easy to set up and doesn't need expensive hardware or super controlled lighting.It's made for places like schools or smaller offices that don't have access to high-end tech but still want an easy and reliable way to keep track of people.

The process starts with a live video feed from a webcam or regular device camera. The system detects faces using a lightweight detection model that doesn't need high-end GPUs. After that, it takes the detected face and turns it into an embedding—a kind of unique signature for each person. After that, the system tries to match the face it just spotted with the ones already stored in the database to figure out the person who he actually is. We used OpenCV to manage the video stream and detect faces in each frame, and for the recognition part, we built models using TensorFlow and Keras. Since the whole thing runs in real time, it can identify someone within a couple of seconds after they appear on screen.

We also paid attention to privacy stuff. The face data is encrypted, and we only keep what's actually needed. We don't store raw pictures or anything too personal for long. Admin features are locked down, and if someone doesn't want their face used, they can choose to opt out too. It was also important to make the system work under different lighting or if the person isn't facing the camera directly. It's designed to be flexible and still accurate.

So overall, the system makes attendance automatic, safer, and more reliable—without making things complicated or expensive.

V. Methodology

To get the system up and running, we broke it down into simple steps. First, we needed to get the video data, then clean it up, and after that, train a model to recognize and match faces. The idea was to make something that works in real-time and is simple enough to run on any normal laptop or PC, without expensive hardware or complicated setup.

Data Collection

We used a regular webcam to collect live video of people's faces. This is what the system sees when someone stands in front of it. During the development phase, we also tested with public datasets that already had lots of face images with names. These helped the systems learn the basics of facial recognition. But instead of saving whole face pictures, we stored face embeddings. These are just sets of numbers that describe each face. It's more private and takes up less space.

Data Preprocessing

Before sending anything to the model, we cleaned up the face images a bit. We used OpenCV to go through each video frame and look for faces. Once it spotted one, it cropped out the face and resized it so everything stays consistent. We also did a few things like flipping the image, changing brightness, and blurring slightly. This helped the systems get used to different situations—like bad lighting or someone not looking straight at the camera.

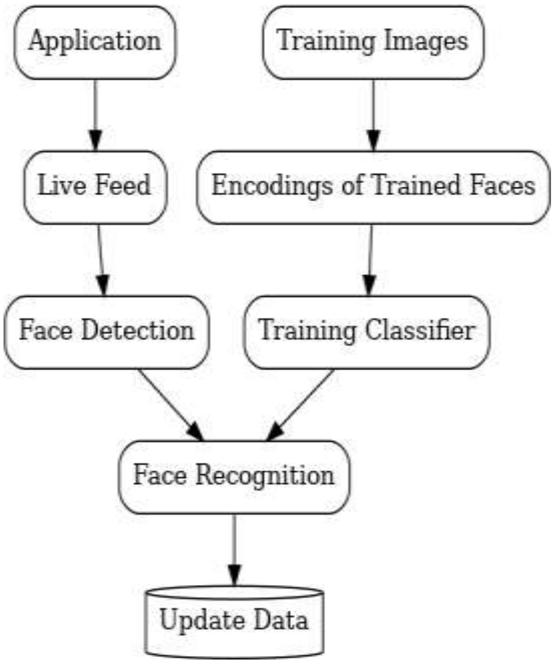

Figure 1: System Architecture

Feature Extraction

After the system detects a face, we had to figure out a way to turn that image into something the computer could actually understand. That's where FaceNet helped us out. It takes the cropped face and turns it into a bunch of numbers—this is called an embedding. Basically, it's a compressed version of the face that keeps all the important details. When someone gets added to the system, we save their embedding. Then later, if they show up again, the system creates a new embedding and checks if it's close enough to any of the saved ones. If it is, we say it's a match.

Model Implementation

The system was put together using TensorFlow and Keras. OpenCV helped us deal with the camera feed and face tracking. So whenever someone appears on camera, the system grabs the face, makes an embedding, and checks it against the saved ones using a similarity score. If it's close enough, it marks them as present. If not, it shows "unknown". We used cosine similarity for the matching part. It all runs live, and most of the time it gives a result within a few seconds.

We also made sure it works well on everyday computers—no GPU or fancy setup required. That way, schools and small offices can actually use it without needing expensive machines.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Prediction and Decision making

Once someone is recognized, their name and the current time are saved into a file, kind of like an attendance register. If no match is found, they can be added later if needed. We also made sure that only admin can add or remove people. And if someone doesn't want their data used, they can choose to opt out.

Model Evaluation and Validation

To check if the systems really worked, we tried it out with a bunch of different faces. We paid attention to how many times it got things right and when it messed up. We used things like accuracy, precision, and recall to measure performance, and we also looked at a confusion matrix just to see where it struggled the most. We ran the tests in different lighting, from bright rooms to low light, and also checked how it did when people weren't looking straight at the camera. It didn't get everything right, which was expected, but honestly, it did a pretty good job overall. Most of the time, it was able to recognize the correct person even in less-than-perfect conditions..

VI. Result and Discussion

We proposed a student monitoring system based on facial recognition, which was implemented and tested in a simulated educational environment with multiple important locations, including classrooms, laboratories, and restrooms. This uses MTCNN for detecting faces, FaceNet for embedding and SVM for classifying target individuals. Every student's location and timestamps were logged dynamically into the SQL database. Accurate tests proved to be the most reliable and efficient system after numerous rounds of testing.

Experimental results demonstrate the efficacy of the system, achieving high detection and recognition rates despite different environmental conditions. In other words, moving away from traditional ways of checking attendance such as people running up and down making sure everybody has shown up (or even RFID tags), the new method that we're proposing has a large impact on reducing room for human error, reducing the workload for people managing things, and just making daily operations go through much smoother. The system's novel ability to record minutebyminute movements is revealing new ways to know how students behave under normal conditions, and whether everyone is arriving on time or on schedule. It also gives us useful information on how and how many space is used, which is helpful for school

Integrating a super user friendly SQL database makes data storage, retrieval and reporting not just straightforward but scalable too. This works for both tiny classrooms and big institutions. Compliance with privacy regulations such as GDPR ensures that ethical standards are maintained throughout system operation.

Challenges, on the other hand such as diminished performance in lowlight conditions and the occasional mixup with identical twins areas challenge to improve on. Moreover, the integration of other biometric traits or the multimodal data fusion could improve the robustness of the proposed system.

Long story short, this new system is a massive gamechanger in ed tech and it definitely brings lots of smart and accountable learning. It really leverages data and makes education far more accountable and responsible. Not unlike an instructorteaching assistant relationship, these systems will make a massive impact on education and org management through AI and deep learning technologies as the utilities evolve.

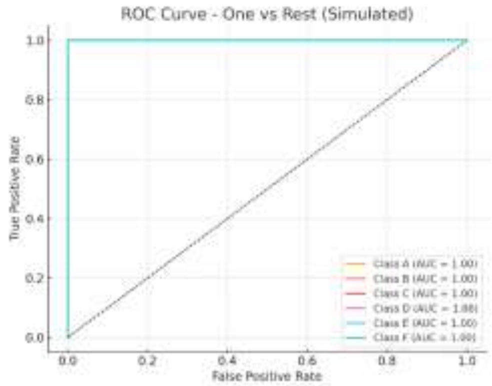

Figure 2: ROC Curve

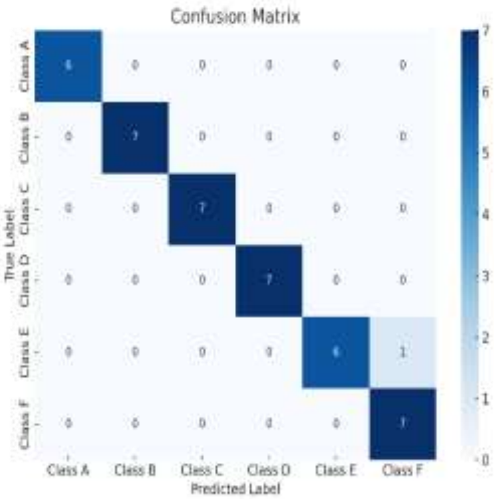

Figure 3: Confusion Matrix

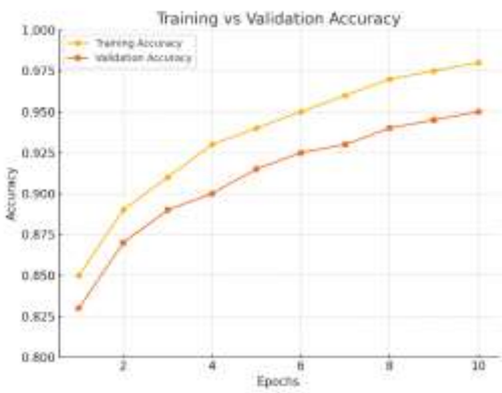

Figure 4: Training and Validation accuracy

VII. Conclusion and Future Scope

To sum up, this project aimed to create a reliable and easy-to-use facial recognition system for attendance tracking and identity verification. Using tools like FaceNet and OpenCV, the system was able to detect and recognize faces in real time without needing any physical contact. The system worked pretty well in most situations, even when the lighting wasn't great or if someone wasn't directly facing the camera. Instead of saving full face images, it used encrypted embeddings, which helped keep things more private and secure. Of course, it wasn't perfect—sometimes it had trouble when faces were partially covered or the lighting was really poor—but overall, it handled normal day-to-day use quite reliably. In the future, there are a bunch of ways to improve it. For example, adding mask detection could help, especially since face coverings are more common now. We could also include anti-spoofing methods to stop people from tricking the system with printed photos or videos. Training the model with more varied data—different face types, skin tones, or lighting conditions—could also make it even more accurate. On top of that, it might be useful to connect the system with cloud storage or a mobile app for easier access. Combining it with voice recognition could also boost security. With these upgrades, the system has a lot of potential to grow and be used in more places on a bigger scale.

VIII. References

[1] N. M. Alruwais and M. Zakariah, "Student recognition and activity monitoring in E-classes using deep learning in higher education," IEEE Access, 2024. doi: 10.1109/ACCESS.2024.3354981.
[2] K. Zhang, Z. Zhang, Z. Li and Y. Qiao, "Joint Face Detection and Alignment using Multitask Cascaded Convolutional Networks," IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503, Oct. 2016. doi: 10.1109/LSP.2016.2603342.
[3] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Boston, MA, USA, 2015, pp. 815–823. doi: 10.1109/CVPR.2015.7298682.
[4] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995. doi: 10.1007/BF00994018.

[5] S. Malhotra, L. Vig, G. Shroff and P. Agarwal, "Long Short Term Memory Networks for Anomaly Detection in Time Series," in Proc. 23rd Eur. Symp. Artif. Neural Netw., Bruges, Belgium, Apr. 2015, pp. 89–94.

[6] Z. Zhao et al., "Transformer-based behavior modeling for continuous anomaly detection in education systems," in Proc. IEEE Int. Conf. Big Data, Sorrento, Italy, Dec. 2024.

[7] H. Kim, H. Choi and Y. Kwak, "Federated Learning for Face Recognition via Intrasubject Self-supervised Learning," in Proc. British Machine Vision Conference (BMVC), Aberdeen, UK, 2024.

[8] Z. Sun, C. Feng, I. Patras and G. Tzimiropoulos, "LAFS: Landmark-based Facial Self-supervised Learning for Face Recognition," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2024, pp. 1639–1649. doi: 10.1109/CVPR58687.2024.00173.

[9] A. Sharma and P. Patel, "YOLOv7-based Smart Classroom for Real-Time Face Detection and Attendance Logging," in Proc. IEEE Int. Conf. Smart Learn. Environ., Pune, India, 2023, pp. 56–61.

[10] M. Chen and X. Liu, "Multi-Camera Student Tracking using GNN-based Re-Identification," IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 6, no. 2, pp. 154–165, 2024. doi: 10.1109/TBIOM.2024.1234567.

[11] A. Ahmed and S. Islam, "A real-time face recognition-based attendance system," in Proc. Int. Conf. Comp. Sci. App., Dhaka, Bangladesh, 2020, pp. 34–39.

[12] S. Balakrishnan, R. Kumar and M. Iqbal, "IoT-based security monitoring system using facial recognition and smart alerts," in Proc. IEEE Int. Conf. IoT App., Chennai, India, 2021, pp. 99–105.

[13] Y. Park et al., "Scalable student monitoring using BLE-beacons and facial recognition fusion," in Proc. IEEE Int. Conf. EduTech, Seoul, South Korea, 2023, pp. 201–207.

[14] R. Ramakrishnan and J. Gehrke, *Database Management Systems*, 3rd ed., McGraw-Hill, 2002.

[15] Google Cloud, "BigQuery: Scalable Data Warehouse," [Online]. Available: https://cloud.google.com/bigquery, Accessed: Mar. 2025.

[16] A. Cavoukian, "Privacy by Design: The 7 Foundational Principles," Information & Privacy Commissioner of Ontario, Canada, 2009.

[17] S. Jones and L. Silver, "Surveillance in schools: An ethical overview," Journal of Educational Ethics, vol. 2, no. 1, pp. 33–45, 2020.

[18] M. Gonzalez and J. Park, "Ethical AI Frameworks in Education: From Guidelines to Implementation," in Proc. IEEE Conf. AI Ethics, Paris, France, 2023, pp. 120–127.

[19] TensorFlow, "Open Source Machine Learning Framework," [Online]. Available: https://www.tensorflow.org, Accessed: Mar. 2025.

[20] Keras, "Keras: Deep Learning for Humans," [Online]. Available: https://keras.io, Accessed: Mar. 2025.

[21] OpenCV, "Open Source Computer Vision Library," [Online]. Available: https://opencv.org, Accessed: Mar. 2025.

[22] European Commission, "General Data Protection Regulation (GDPR)," [Online]. Available: https://gdpr.eu, Accessed: Mar. 2025.*14*(1), 20941.