

Student Attendance Using CCTV and Computer Vision: A Systematic Review of Existing Approaches, Challenges and Research Gaps

Asst. Prof. Paritosh Nagarnaik¹, Arhant Kamble², Ujwal Thakre³,

Anshul Kothare⁴, Mayank Bisnewar⁵

Department of Artificial Intelligence¹⁻⁵

Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

paritoshnagarnaik91@gmail.com¹, arhantqwe@gmail.com², ujwalthakre35@gmail.com³,

kothareanshul01@gmail.com⁴, mayankbisnewar@gmail.com⁵

Abstract: This paper presents an automated student attendance system integrating CCTV video, YOLO-based student detection, OpenCV image processing, and FaceNet-based facial recognition for classroom environments. The proposed system is designed to operate at an approximate distance of 8–10 meters, enabling attendance to be recorded without requiring students to approach a dedicated device. CCTV video frames are processed to detect students and extract facial regions for identification. YOLO is employed for student detection, while OpenCV performs frame extraction, resizing, preprocessing, and camera-stream management. FaceNet generates facial embeddings that are compared with registered student facial data to verify identities. To address low-resolution facial images caused by distance, an optional super-resolution stage may be incorporated before recognition. Five facial images from each registered student are used to establish reference data for identification. Furthermore, multi-frame and confidence-based verification can reduce false recognition caused by temporary blur, occlusion, or inaccurate detections. Once a student's identity is sufficiently confirmed, the system automatically records the attendance status, student name, date, and time. Dedicated Admin and Teacher panels provide controlled management of student information and attendance records. The system will be evaluated at different distances and classroom conditions using recognition accuracy, precision, recall, false-recognition rate, processing time, FPS, and attendance-recording success. The proposed framework aims to provide a practical, contactless, scalable, and reliable solution for automated educational attendance management.

Keywords: Attendance System; Computer Vision; Face Recognition; OpenCV; Student Monitoring; YOLO

I. INTRODUCTION

Attendance management is essential for monitoring student participation and maintaining accurate academic records; however, conventional methods such as manual roll calls and paper registers are time-consuming, error-prone, and vulnerable to proxy attendance. Recent advances in artificial intelligence and computer vision offer efficient, contactless alternatives, yet many existing facial-recognition systems perform primarily under controlled, close-range conditions. In practical classrooms, variations in distance, facial scale, illumination, pose, occlusion, motion blur, and image resolution can significantly affect recognition reliability. This research proposes a **CCTV-based automated attendance system** integrating YOLO-based student detection, OpenCV image processing, and FaceNet-based facial recognition for an approximate **8–10 meter** operating range. Optional image enhancement or super-resolution will be employed for low-quality facial regions, while confidence-based and multi-frame verification will improve



identification reliability and minimize duplicate or false attendance. The system will be evaluated under realistic classroom conditions using recognition accuracy, precision, recall, false-recognition rate, processing time, FPS, and attendance-recording success rate.

II. LITERATURE REVIEW

[1] **Sani et al. (2025)**. Sani et al. proposed a real-time attendance system combining YOLOv11, FaceNet, and super-resolution models to address low-quality facial images captured from longer distances. ESRGAN, Real-ESRGAN, and GFPGAN were investigated for facial enhancement. The study is highly relevant to the proposed work because it directly addresses the relationship between distance and facial-recognition quality. However, further evaluation under a controlled **8–10 m classroom environment** remains valuable.

[2] **Sawhney et al. (2019)**. Sawhney et al. developed a real-time smart attendance system based on facial-recognition techniques. Their work demonstrates the feasibility of automating attendance while reducing manual effort and improving record management. The study establishes an important foundation for face-recognition-based attendance; however, practical challenges associated with distant faces, classroom-scale monitoring, illumination variation, and occlusion require further investigation. These limitations motivate the proposed CCTV-oriented approach.

[3] **Zhang et al. (2024)**. Zhang et al. investigated low-quality face recognition using a texture-guided transfer-learning approach. Their method employs knowledge distillation and gradient-texture information to reduce discrepancies between high- and low-quality facial representations. An attention-based variant further emphasizes informative facial regions. Experiments on TinyFace and QMUL-SurFace demonstrated improved recognition performance. This work is particularly relevant to CCTV attendance, where distant students may produce low-resolution facial images.

[4] **Wang et al. (2018)**. Wang et al. proposed ESRGAN to improve single-image super-resolution by enhancing network architecture, perceptual loss, and adversarial learning. The method produces more realistic high-resolution images than earlier approaches such as SRGAN. For the proposed attendance system, ESRGAN is relevant because distant CCTV images may contain insufficient facial detail for reliable recognition. Nevertheless, improved visual quality does not necessarily guarantee improved identity recognition, making recognition-level evaluation essential.

[5] **Wang et al. (2021)**. Wang et al. introduced GFP-GAN for real-world blind face restoration under complex degradation conditions. The method incorporates a pretrained generative facial prior to restore facial details and enhance colour in a single forward pass. Its ability to process severely degraded facial images makes it potentially useful for distant CCTV applications. However, generated facial details must be carefully evaluated because restoration should preserve identity rather than introduce misleading information.

[6] **Schroff et al. (2015)**. Schroff et al. introduced FaceNet, which directly maps facial images into a compact Euclidean embedding space. Facial similarity can consequently be determined by measuring distances between embeddings. The approach provides an efficient foundation for face verification, recognition, and clustering. Its embedding-based representation is suitable for attendance systems because registered students can be represented through facial vectors and subsequently compared with faces detected from classroom video.

[7] **Rao (2022)**. Rao proposed AttenFace, an automated attendance system that identifies students from classroom snapshots captured throughout a lecture. Instead of relying on a single recognition event, attendance is granted according to a student's presence across multiple snapshots. This temporal strategy reduces the possibility of incorrectly marking a student based on brief appearance. The approach is particularly relevant to the proposed system's multi-frame attendance-decision strategy.

[8] **Bhattacharya et al. (2018)**. Bhattacharya et al. proposed the Smart Attendance Monitoring System (SAMS), a face-recognition-based solution developed specifically for classroom attendance. The study addresses the limitations of conventional attendance practices, particularly their time consumption and reliability concerns. Its classroom-specific focus makes it directly relevant to the proposed research. However, the proposed 8–10 m CCTV configuration introduces additional distance and low-resolution challenges beyond conventional classroom face-recognition scenarios.



Sr. No.	Title	Author	Description	Conclusion
[1]	Improving Real-Time Attendance System Based on YOLO and FaceNet Using Super Resolution Models	M. F. Sani, R. F. Rachmadi, Endroyono	Addresses low-quality facial images captured from long distances in real-time attendance systems.	YOLOv11 + FaceNet + ESRGAN/Real ESRGAN/GFPGAN. GFPGAN produced realistic enhancement, while the combination of YOLOv11, GFPGAN and FaceNet improved distant-face recognition.
[2]	Real-Time Smart Attendance System using Face Recognition Techniques	S. Sawhney, K. Kacker, S. Jain, S. N. Singh, R. Garg	Develops an automated attendance system to reduce the burden and time associated with manual attendance.	LBP + PCA + SVM. The study demonstrated the feasibility of face-recognition-based automated attendance with reported recognition performance.
[3]	Texture-Guided Transfer Learning for Low-Quality Face Recognition	M. Zhang, R. Liu, D. Deguchi, H. Murase	Focuses on the difficult problem of recognizing faces from low-quality surveillance images.	Texture-Guided Transfer Learning + Knowledge Distillation + Attention. The method improved low-quality face recognition, particularly on TinyFace and QMUL-SurFace.
[4]	ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks	X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, C. C. Loy	Improves single-image super-resolution by generating more realistic high-resolution image details.	ESRGAN. Uses improved architecture, perceptual loss and adversarial learning to enhance low-resolution images.
[5]	Towards Real-World Blind Face Restoration with Generative Facial Prior	X. Wang, Y. Li, H. Zhang, Y. Shan	Addresses face restoration when images contain unknown degradation such as blur, noise, compression and low resolution.	GFP-GAN. Uses a pretrained generative facial prior to restore degraded facial images while improving perceptual quality.
[6]	FaceNet: A Unified Embedding for Face Recognition and Clustering	F. Schroff, D. Kalenichenko, J. Philbin	Introduces a method for representing faces as compact numerical embeddings that can be compared for identity verification.	FaceNet + Triplet Loss. Learns embeddings in which faces of the same person are close while different identities are separated.
[7]	AttenFace: A Real Time Attendance System Using Face Recognition	A. Rao	Uses classroom snapshots throughout a lecture instead of relying on continuous recognition from every video frame.	Face Recognition + Multi-Snapshot Attendance. Students are marked present when they appear in a sufficient number of snapshots, reducing unnecessary computation and transient recognition errors.
[8]	Smart Attendance Monitoring System (SAMS): A Face Recognition Based	S. Bhattacharya et al.	Presents a face-recognition-based attendance system specifically designed	Face Recognition + Classroom Attendance Monitoring. Demonstrates the feasibility of automating classroom attendance while reducing manual attendance effort.



Sr. No.	Title	Author	Description	Conclusion
	Attendance System for Classroom Environment		for classroom environments.	

III. RESEARCH METHODOLOGY -

The proposed system employs a structured computer-vision pipeline for automated attendance using CCTV video. During registration, each student will receive a dedicated profile, and **five facial images** will be captured to establish reliable reference data for recognition. During attendance, a strategically positioned CCTV camera will capture classroom video within an approximate **8–10 meter** range. OpenCV will handle frame extraction and preprocessing, while YOLO will detect students and isolate facial regions. These faces will be represented using **FaceNet embeddings** and compared with registered facial data. Where necessary, super-resolution or image enhancement will be applied to improve low-quality facial regions. **Confidence-based and multi-frame verification** will be used to strengthen identification and reduce false or duplicate attendance. Once a student's identity is confirmed, the system will automatically record the student's name, attendance status, date, and time. Separate **Admin and Teacher panels** will facilitate centralized management, including student registration, facial data, attendance sessions, monitoring, and records. This architecture is intended to provide a scalable, efficient, and reliable automated attendance solution.

IV. RESULTS AND DISCUSSION

The proposed system addresses key limitations identified in existing research, including **distant face recognition, low-resolution images, illumination variation, partial occlusion, and unreliable single-frame identification**. Unlike systems evaluated primarily under controlled conditions, the proposed framework will be tested in a realistic classroom environment at **8–10 meters** from the CCTV camera. The system will employ a multi-stage pipeline comprising **student detection, facial-region extraction, preprocessing, optional super-resolution, facial embedding, and confidence-based verification**. Five facial images from each registered student will establish individual reference representations. **Multi-frame confirmation** will be used to strengthen identity decisions and reduce errors caused by temporary blur, occlusion, or incorrect detections.

Performance will be evaluated at **8 m, 9 m, and 10 m** using recognition accuracy, precision, recall, false-recognition rate, missed detections, processing time, FPS, and attendance-recording success. The effect of image enhancement on distant-face recognition will also be examined.

The primary research gap is the limited **distance-specific and classroom-oriented evaluation** of automated attendance systems. By integrating real-time detection, facial embeddings, multi-frame verification, optional enhancement, and dedicated **Admin and Teacher panels**, the proposed system aims to provide a practical and robust attendance framework. Final results will be based exclusively on experimental measurements.

V. CONCLUSION

The proposed CCTV-based automated student attendance system aims to provide a reliable, contactless, and efficient alternative to conventional attendance methods. By integrating YOLO-based student detection, OpenCV preprocessing, FaceNet-based facial recognition, and optional super-resolution, the system is designed to address challenges associated with distant and low-quality facial images. The use of five facial images per student provides reference data for identity verification, while multi-frame and confidence-based recognition can improve the reliability of attendance decisions. Furthermore, dedicated Admin and Teacher panels enable effective management of student information, facial data, attendance sessions, and records. The proposed system will be evaluated under realistic classroom conditions, particularly across an **8–10 meter operating range**, using appropriate performance metrics. Overall, the research seeks



to establish a scalable and practical attendance framework that reduces manual effort, minimizes attendance-related errors, and demonstrates the applicability of computer vision for automated classroom attendance management.

REFERENCES

- [1] Sani, M. F., Rachmadi, R. F., & Endroyono. (2025). Improving real-time attendance system based on YOLO and FaceNet using super resolution models. *2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI)*, 113–118.
- [2] Sawhney, S., Kacker, K., Jain, S., Singh, S. N., & Garg, R. (2019). Real-time smart attendance system using face recognition techniques. *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 522–525.
- [3] Zhang, M., Liu, R., Deguchi, D., & Murase, H. (2024). Texture-guided transfer learning for low-quality face recognition. *IEEE Transactions on Image Processing*, 33, 95–107.
- [4] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (2018). ESRGAN: Enhanced super-resolution generative adversarial networks. *European Conference on Computer Vision Workshops*, 63–79.
- [5] Wang, X., Li, Y., Zhang, H., & Shan, Y. (2021). Towards real-world blind face restoration with generative facial prior. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9164–9174.
- [6] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823.
- [7] Rao, A. (2022). AttenFace: A real time attendance system using face recognition. *2022 IEEE International Conference on Contemporary Computing and Communications (InC4)*.
- [8] Bhattacharya, S., Nainala, G. S., Das, P., & Routray, A. (2018). Smart attendance monitoring system (SAMS): A face recognition based attendance system for classroom environment. *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*.

