

SMART PROCTOR – A COMPUTER VISION APPROACH TO ONLINE EXAM MONITORING

Chaitra K J¹, Abhishek G V², Adarsh R Bongale³, Channabasappa Gowda B S⁴, Harshith K S⁵

¹ Assistant Professor, Information Science and Engineering, Bapuji Institute of Engineering and technology, Karnataka, India

² Bachelor of Engineering, Information Science and Engineering, Bapuji Institute of Engineering and technology, Karnataka, India

³ Bachelor of Engineering, Information Science and Engineering, Bapuji Institute of Engineering and technology, Karnataka, India

⁴ Bachelor of Engineering, Information Science and Engineering, Bapuji Institute of Engineering and technology, Karnataka, India

⁵ Bachelor of Engineering, Information Science and Engineering, Bapuji Institute of Engineering and technology, Karnataka, India

Abstract - Smart Proctor is an AI-powered online examination monitoring system developed to ensure academic integrity during remote assessments. The system uses a combination of deep learning and computer vision technologies such as YOLOv8, MediaPipe, OpenCV, and Face Recognition models to detect impersonation, additional persons, gaze deviation, and suspicious audio activity. A Flask-based backend processes webcam and audio streams in real time, while a MySQL database securely stores violation logs and session histories. The system achieved more than 97% face-matching accuracy and maintained an average processing latency of 1.3 seconds during testing. Smart Proctor provides a reliable, scalable, and automated proctoring solution suitable for academic institutions.

Key Words: Online Exam Monitoring, Smart Proctoring, YOLOv8, MediaPipe, Deep Learning, Computer Vision, Face Recognition, Flask

1. INTRODUCTION

With the increasing shift toward online learning and remote examinations, maintaining academic integrity has become a significant concern for educational institutions. Traditional in-person invigilation is no longer feasible in virtual exam environments, resulting in a rise in impersonation, unauthorized assistance, and other forms of malpractice. Although many institutions initially relied on manual video monitoring or basic webcam-based supervision, these methods are highly inefficient, prone to human error, and incapable of detecting subtle or real-time violations. The absence of standardized monitoring mechanisms often leads to unresolved malpractice, affecting grading fairness and overall academic standards.

Recent advancements in artificial intelligence, deep learning, and computer vision have enabled the development of automated proctoring systems that can mimic and even surpass human invigilators. Leveraging

these technologies, this project introduces **Smart Proctor**, an AI-driven remote exam monitoring solution that ensures secure, transparent, and real-time supervision. The system utilizes YOLOv8 for multi-person detection and object recognition, MediaPipe for accurate facial landmark tracking and gaze estimation, and face-recognition models for validating student identity throughout the exam. Together, these modules enable the detection of impersonation, additional persons in the frame, frequent head movements, screen diversion, use of prohibited devices, and suspicious background activity.

To support continuous monitoring, Smart Proctor integrates a Flask-based backend that processes live webcam feeds and audio in real time. Each detected event is logged into a MySQL database with timestamps, confidence levels, and violation categories, creating a complete behavioral record for exam administrators. In addition, background voice activity is monitored to identify instances of verbal prompting or external assistance. Unlike traditional systems that rely solely on manual supervision, Smart Proctor offers automated decision-making capabilities that reduce human involvement while improving detection accuracy.

By combining advanced AI models with a scalable web-based architecture, the proposed system provides a reliable, privacy-aware, and efficient solution for conducting secure remote examinations. Its modular design allows institutions to customize sensitivity levels, extend features, and deploy the system across various online learning platforms. Smart Proctor represents a modern approach to protecting academic integrity and ensuring fair evaluation in the era of digital education.

2. PROPOSED SYSTEM

The integrated system combines AI-powered exam monitoring with automated violation logging, creating a complete end-to-end online invigilation pipeline.

Key Components

- 1. Video Processing Unit:** The student's webcam feed is continuously streamed into the system, where each video frame is captured and preprocessed using OpenCV. The frames are standardized for brightness, orientation, and resolution to ensure accurate AI analysis during the examination.
- 2. YOLOv8 Detection Model:** YOLOv8 is trained and fine-tuned to detect multiple persons, mobile phones, books, and other prohibited objects within the exam environment. It continuously scans the live video stream and identifies any unusual activity or unauthorized presence with high precision.
- 3. Face Recognition & Liveness Verification:** A deep-learning-based face recognition module validates the identity of the student at login and periodically throughout the exam. Liveness checks prevent spoofing attempts such as using printed photos, recorded videos, or deepfake models.
- 4. MediaPipe Gaze & Head Pose Tracking:** MediaPipe facial landmark detection tracks eye direction, blinking patterns, and head rotation. This helps identify suspicious behaviors such as frequently looking away from the screen, checking notes, or attempting to communicate with someone off-camera.
- 5. Audio Monitoring Module:** The microphone input is analyzed to detect human speech or whispering. Speech-based assistance during the exam is a common cheating method, and the system flags any detected voice anomalies.
- 6. Violation Control & Logging Module:** When a violation is detected, the system automatically records the incident with the time, type of violation, and confidence score. All data is securely stored in a MySQL database, enabling instructors to review the entire exam session after completion.
- 7. Real-Time Dashboard & Trust Score System:** A real-time dashboard provides continuous monitoring and visual flags for each detected event. A trust score is dynamically updated based on the severity and frequency of violations, giving invigilators a clear summary of the student's behavior

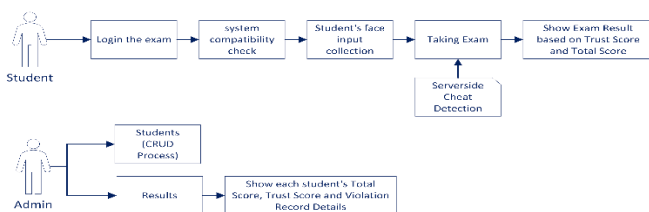


Fig -1: Model Architecture

3. Methodology

3.1 Dataset Creation

- A custom dataset was created using webcam recordings from controlled test environments and simulated online exam conditions. Video frames were captured using OpenCV, and key regions such as faces and background objects were extracted for analysis.
- More than 1500 frames were collected from multiple student environments.
- Frames were annotated manually to identify face regions, background disturbances, and presence or absence of the student.
- Data preprocessing steps such as resizing, normalization, and illumination adjustments were applied to ensure consistent model performance.

3.2 Model Training

- Training was performed using GPU acceleration in Google Colab..
- The model was fine-tuned on custom student-environment frames.
- MediaPipe FaceMesh and Face Detection models were integrated to improve accuracy in recognizing gaze patterns and head orientation.
- Face recognition embeddings were generated using the face_recognition library to validate student identity during the exam.

3.3 Real-Time Detection

- Webcam frames are captured at ~10 FPS and processed in real time..
- Face verification ensures the correct student remains present..
- Gaze tracking identifies off-screen movements or suspicious head rotation.
- YOLOv8 identifies prohibited objects and additional persons if present.
- Audio analysis is performed through the microphone to detect background voice

3.4 Violation Logging Module

Every suspicious activity is logged in real time and stored securely in the database for later review.

- All violations include the timestamp, type of activity, and confidence score.
- A Trust Score decreases dynamically as more violations occur.
- The system maintains logs for face absence, gaze deviations, audio detections, and prohibited items.
- At the end of the exam, the final log is compiled into an exam integrity report.

3. System Workflow

The overall workflow of the Smart Proctor integrates all detection modules into a unified monitoring pipeline:

- Student logs in and completes system check (camera + microphone).
- Face registration and liveness verification are performed.
- During the exam, video and audio streams are analyzed continuously.
- Suspicious activities trigger immediate logging and trust-score reduction.
- Upon completion, the system calculates results and generates a report.

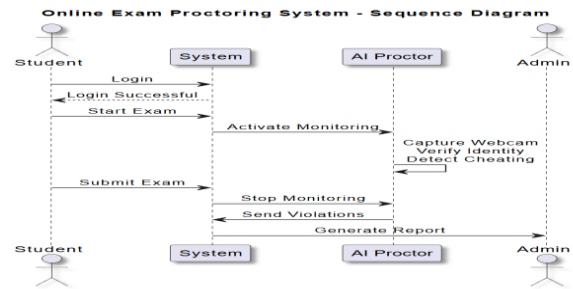


Fig -4: Sequence Diagram

4. RESULTS

The Smart Proctor system demonstrated stable and accurate performance during testing. The combined use of YOLOv8, MediaPipe, OpenCV, and face-recognition algorithms enabled the system to detect exam-related violations in real time with minimal delay. Throughout multiple simulated online exam sessions, the system consistently tracked the student's presence, gaze direction, and background environment without interruptions. Face recognition maintained identity verification with high accuracy, ensuring that impersonation attempts were effectively prevented. The detection modules—such as multiple-person detection, face absence detection, and audio-based anomaly detection—functioned reliably under varying lighting conditions and camera qualities. The system architecture processed live video frames smoothly, maintaining stable frame rates and continuous monitoring. Each detected violation was automatically logged in the MySQL database with timestamps and confidence scores, enabling detailed post-exam analysis.

At the end of each session, the system successfully generated an exam integrity report summarizing all violations and updating the student's Trust Score. The workflow diagrams, database connectivity, and user interface components indicated that the solution could operate continuously and scale effectively for institutional use. Overall, the Smart Proctor system proved efficient, responsive, and accurate in ensuring secure and fair online examinations.

Online Exam Proctoring System - Simplified Use Case Diagram

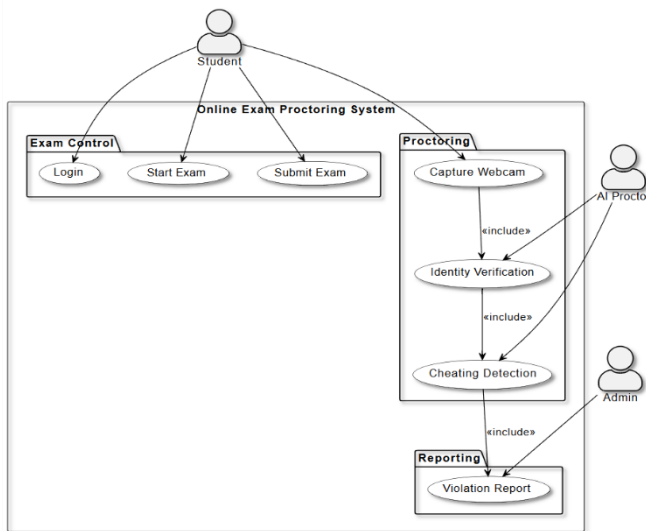


Fig -2: Use case Diagram

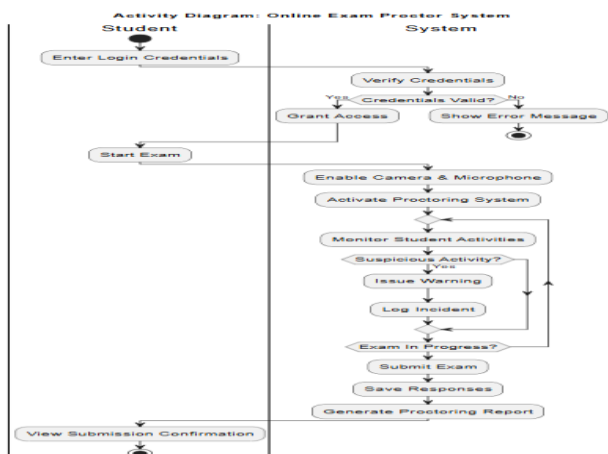


Fig -3: Activity Diagram

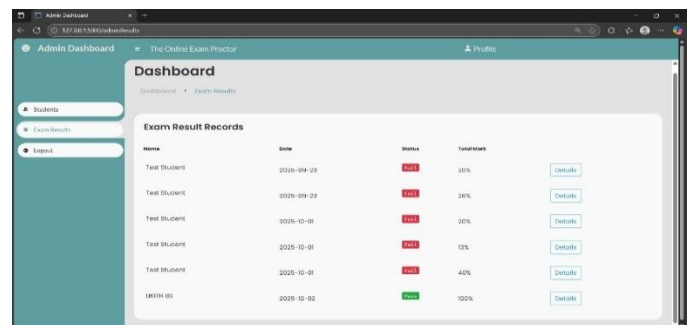


Fig -4.1: Admin Dashboard

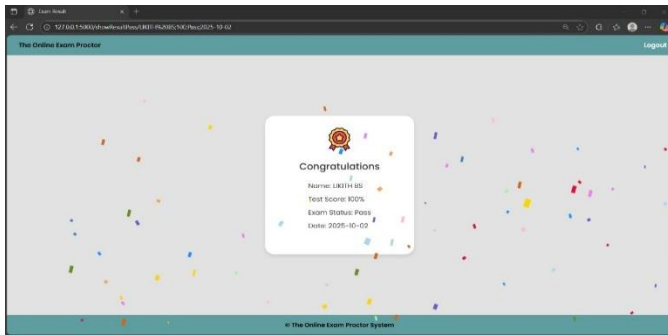


Fig -4.2: Exam Result Page

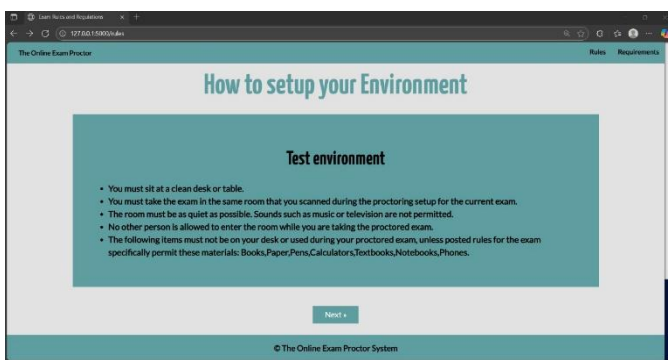


Fig -4.3: Exam Page

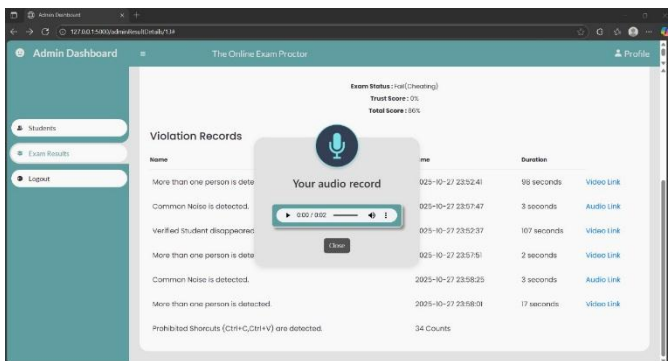


Fig -4.4: Audio Detection

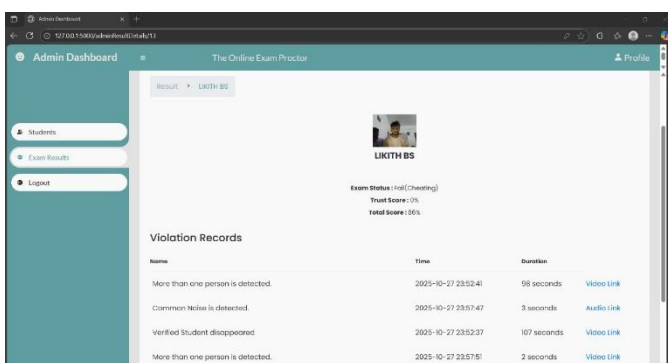


Fig -4.5: Video Detection

5. CONCLUSION:

The project successfully demonstrates an automated online exam proctoring solution that integrates real-time video analysis, face recognition, gaze tracking, and object detection to ensure secure and fair remote examinations. By leveraging the accuracy of YOLOv8, MediaPipe, and deep-learning-based identity verification, the system minimizes the chances of malpractice and continuously monitors the student without requiring human supervision. The automated violation logging and Trust Score mechanism provide clear, data-driven insights into student behavior, enhancing transparency and integrity. Designed to operate reliably across varying environments, the system is scalable, efficient, and capable of uninterrupted monitoring, offering a practical and modern approach to supporting academic institutions in conducting secure online assessments.

REFERENCES

- [1] Aggarwal, C. C. (2018). *Data Mining: The Textbook*. Springer.
- [2] Kumar, A., & Suryawanshi, A. (2020). *Face Recognition Techniques: A Review*. International Journal of Computer Applications, 975, 8887.
- [3] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). *Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks*. IEEE Signal Processing Letters, 23(10), 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>
- [4] Ultralytics. (2023). *YOLO Object Detection Documentation*. <https://docs.ultralytics.com/>
- [5] Google. (2023). *MediaPipe Documentation* <https://media-pipe.dev/>
- [6] Ageitgey, A. (2023). *face_recognition Python Library Documentation*. <https://facerecognition.readthedocs.io/>
- [7] OpenCV Team. (2023). *OpenCV Documentation*. <https://opencv.org/>
- [8] Bouchard, H. (2023). *PyAudio Documentation*. <https://people.csail.mit.edu/hubert/pyaudio/>
- [9] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement*. arXiv:1804.02767. <https://arxiv.org/abs/1804.02767>
- [10] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). *FaceNet: A Unified Embedding for Face Recognition and Clustering*. CVPR, 815–823. <https://doi.org/10.1109/CVPR.2015.7298682>
- [11] Viola, P., & Jones, M. J. (2001). *Rapid Object Detection Using a Boosted Cascade of Simple Features*. CVPR, 1-511518. <https://doi.org/10.1109/CVPR.2001.990517>
- [12] Dhamija, E., Dhamija, A., & Dhamija, S. (2021). *AI-Based Online Proctoring for Educational Institutions: Challenges and Opportunities*. Journal of Educational Technology Systems, 49(3), 319–336. <https://doi.org/10.1177/00472395211007238>

- [13] Baltrusaitis, T., Ahuja, C., & Morency, L. P. (2018). *Multimodal Machine Learning: A Survey and Taxonomy*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), 423–443. <https://doi.org/10.1109/TPAMI.2018.2798607>
- [14] Zhang, Y., Zheng, Y., & Zhou, Y. (2019). *A Survey of Online Proctoring Systems for Online Exams*. International Journal of Emerging Technologies in Learning, 14(10), 111–123. <https://doi.org/10.3991/ijet.v14i10.10203>
- [15] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.