International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 06 June-2026 | Impact Factor: 3.5

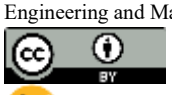
EDU Proctorx: AI Powered Remote Assessment Monitor

Gulam Mahabub Subani[1] , Kesari Karthik[2] , CH.Sai Krishna[3] , Md. Tousif[4]

[1]Assistant Professor Department of Computer Science and Engineering (AI & ML) CMR Technical Campus (UGC Autonomous), Kandlakoya, Hyderabad, Telangana.
[2,3,4]UG Scholar Department of Computer Science and Engineering (AI & ML) CMR Technical Campus (UGC Autonomous), Kandlakoya, Hyderabad, Telangana.

[1]gulammahabubsubhani.aiml@cmrtc.ac.in,[2]kesarikarthik147@gmail.com,[3]chandavolusaikrishna2005@gmail.com,[3]nadmantwasifma@gmail.com

How to Cite this Article:
Karthik, K., Krishna, C. & Tousif, M. (2026). EDU Proctorx: AI Powered Remote Assessment Monitor. International Journal of Creative and Open Research in Engineering and Management, 02(6). https://doi.org/10.55041/ijcope.v2i6.238



https://doi.org/ 10.55041/ijcope.v2i6.238

ABSTRACT:

The rapid proliferation of online exams has led to a significant demand for reliable systems that uphold academic integrity, fairness, and security in remote assessment settings, and since monitoring students without physical supervision is challenging and conventional methods are often inadequate, this paper proposes the AI Powered Remote Assessment Monitor, a multimodal proctoring framework aimed at ensuring a secure and unbiased examination environment by integrating computer vision, audio processing, and behavioral modeling techniques to continuously monitor candidate activities during an examination, with identity checks based on facial recognition and liveness detection, supplemented by gaze tracking, head pose estimation, object detection, and sound classification methods to detect suspicious behaviors, thereby leveraging a multimodal analysis of signals to yield more accurate insights into candidate activities. First, unlike traditional rule-based or entirely manual methods, the framework emphasizes automated monitoring through modular AI elements and a weighted decision fusion approach that integrates data from various sources to reach balanced and contextually aware decisions instead of relying on isolated alerts; also, to ensure responsible utilization, the platform incorporates secure data handling protocols, including encrypted logging and privacy-conscious monitoring; the solution is built to accommodate both live and recorded examinations, and it can be implemented by educational institutions, certification bodies, and training organizations (ibid).

Keywords: Artificial Intelligence, Computer Vision, Machine Learning, Facial Recognition, Behaviour Analysis, Ethical AI, Online Proctoring, Secure Assessments.

. Introduction

The transformation of examination processes has been brought about by the rising popularity of online and hybrid modes of education, which, despite the advantages of remote assessments such as flexibility, accessibility, and scalability, have precipitated critical issues concerning academic honesty and supervision, posing significant challenges in ensuring that students adhere to examination rules even in the absence of physical invigilators, as conventional proctoring methods, predominantly reliant on human supervisors, prove inefficient and costly on a large scale,

hindered further by fatigue, delayed response to incidents, and incurred subjectivity, thereby rendering these methods incompetent for digital examinations on a large scale, whereas many of the available online proctoring mechanisms involve simple procedures like browser locking and basic motion detection that, although useful in providing a basic level of analysis, result in false alarms and lack contextual competence in analyzing candidate behavior, with singlesource analysis, such as video alone, unable to identify complex and elaborate forms of misconduct. To that end, the system uses facial recognition authentication combined with liveness checks to verify the candidate's identity before continuously monitoring gaze, head pose, objects in the room, and audio cues, with an important framework contribution being a weighted decision fusion strategy that aggregates multimodal information to generate an Exam Integrity Score instead of binary flags, allowing for more accurate and context-aware decisions, with additional system features including strong encryption, privacy protection, and ethical AI use that promote transparency and fairness.

The proposed solution is meant for both synchronous and asynchronous examinations and can be adopted across educational institutions, certification, and training organizations to allow secure and trustworthy remote assessments.

2. LITERATURE REVIEW

2.1 Automated Proctoring Frameworks

With the increased adoption of online learning, the practice of administering examinations from a distance has similarly gained prominence. Saha et al. [1] proposed an AIbased intelligent proctoring system integrating facial recognition and motion tracking to improve exam security and fairness. Similarly, Sridhar and Rajshekhar

[7] developed an AIintegrated proctoring framework capable of continuous candidate authentication and anomaly detection. Niharika and Nayak [10] introduced a lightweight AIbased proctoring system emphasizing facial authentication and activity logging The architectural evolution of such systems is

reflected in the modular design presented in Fig.1

2.2 Behaviour Monitoring Techniques This section focuses more on the behaviour patterns of the student rather than taking every single

movement and making countermeasures, because effectively, there are some movements which may not really be suspicious and the intelligent system should be able to consider that fact. To this effect, sometimes, Eye gaze tracking, facial analysis, and posture estimation are widely used techniques to assess attention levels during examinations [6]. Sequential learning models such as Long Short-Term Memory (LSTM) networks are employed to analyze behavioral trends over time instead of single-frame events [5].

2.3 Computer Vision Applications Computer vision plays an important role in many automated proctoring solutions. Videobased monitoring allows the system to observe the candidate's surroundings and detect the presence of unauthorized objects or additional people. For example, detection models can identify items such as mobile phones, books, or secondary screens that may be used to gain unfair assistance.

At the same time, facial recognition methods are widely used to confirm that the correct candidate remains present throughout the exam. Continuous identity checks help prevent impersonation or proxy test-taking.Hence, computer vision plays a paramount role in many automated proctoring solutions

2.4 Ethical and Privacy Considerations Although AI-based monitoring improves exam security, it also raises important ethical and privacy concerns. Excessive surveillance can make students uncomfortable and may increase anxiety during tests. Therefore, recent studies emphasize the need to balance security with fairness and transparency.

Researchers suggest incorporating privacyaware practices such as minimal data storage, encryption,

and clear consent mechanisms. Providing understandable explanations for system decisions is also important to build trust among users. A responsible proctoring system should protect academic integrity while still respecting the rights and well-being of students.

3. METHODOLOGY

This section describes the systematic methodology adopted for designing and implementing the AI Powered Remote Assessment Monitor, a multimodal proctoring framework intended to ensure secure, fair, and reliable online examinations. The proposed approach follows a modular systemdevelopment pipeline that integrates biometric verification, visual and audio monitoring, temporal behaviour analysis, and decision fusion under strict security and ethical constraints.

3.1 System Planning and Architecture Design

3.1.1 Requirement Analysis An initial requirement analysis was conducted to identify the functional and non-functional objectives of an automated proctoring system suitable for realworld academic assessments. The primary requirements include:

Identity verification: Ensuring that the registered candidate is the individual taking the examination through biometric authentication.

Continuous monitoring: Providing uninterrupted supervision of the candidate's visual and acoustic environment throughout the exam session.

Behavioural interpretation: Differentiating between natural test-taking behaviour and potential misconduct using temporal context.

Low latency and scalability: Supporting realtime inference on heterogeneous student devices without requiring high-end hardware. Data security and privacy: Protecting sensitive multimedia data through encryption and ethical data-handling practices.

Faculty feedback and simulated exam scenarios were incorporated to refine these requirements and align the system with institutional policies.

3.1.2 System Architecture The proposed system follows a modular hybrid architecture, where data acquisition occurs on the client device while computationally intensive inference is performed on a secure server or optimized local model. The architecture comprises the following components: **a)**Pre-Exam Authentication Module

b)Computer Vision Monitoring Engine

c) Audio Event Analysis Module

d)Behavioral Pattern Analyzer

e) Multimodal Decision Fusion Engine

f) Alert and Reporting Module

g)Secure Logging and Encryption Layer

This modular design enhances maintainability; scalability, and adaptability in synchronous and asynchronous examination formats.

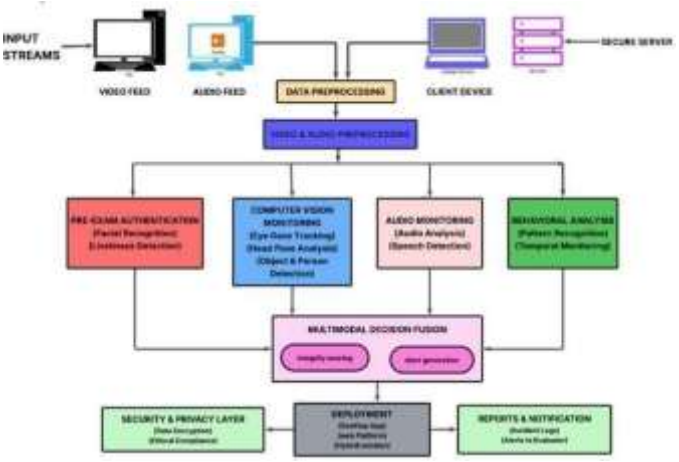


FIG.1: AI POWERED REMOTE ASSESSMENT MONITOR ARCHITECTURE

3.2 Dateset Preparation and Preprocessing

3.2.1 Visual Dataset Collection

Visual data consist of face images, gaze directions, head orientations, object presence, and background activity. Data were collected from a combination of: Publicly available face and gaze datasets

Controlled recordings simulating real examination conditions

The dataset includes variations in lighting, camera resolution, facial orientation, and background environments to improve model robustness and reduce demographic bias.

3.2.2 Audio Dataset Collection Audio samples include: **a)**Human speech

b)Background conversations

c) Device alerts

d)Ambient noise

Pre processing steps include noise filtering, segmentation into short time windows, amplitude normalization, and conversion into Melspectrometer representations.

3.2.3 Data Annotation

All video frames and audio segments were manually annotated using predefined labels such as:

a) Face present / absent

b) Gaze on-screen / off-screen

c) Additional person detected

d) Suspicious audio event

These annotations support supervised training and validation of the individual AI models.

3.3 Identity Verification and Liveness Detection

3.3.1 Facial Recognition

A lightweight convolution neural network (e.g., MobileFaceNet) is used to extract facial embedding. Identity verification follows a fourstage pipeline: **a)** Face detection

b) Feature extraction

c) Embedding vector generation

d) Similarity comparison with stored templates Cosine similarity is used to verify identity against a predefined threshold.

3.4.1 Gaze and Head Pose Estimation

Eye gaze direction and head pose (pitch, yaw, roll) are estimated continuously. Repeated deviations beyond empirically defined thresholds are marked as potential anomalies.

3.4.2 Person and Object Detection

A real-time object detection model (YOLOv5) is employed to identify: Additional individuals

- Unauthorized objects such as smartphones, books, or secondary screens

Detected violations are timestamped and forwarded to the decision fusion module.

3.5 Audio-Based Surveillance

Audio streams are segmented into short frames and transformed into Mel-spectrograms. A CNNbased classifier categorizes sounds into:

An overall Exam Integrity Score (EIS) is computed as:

Temporal modelling reduces false positives caused by isolated, non-intentional actions.

3.6.2 Decision Fusion Mechanism

Let Vi, Ai and Bi represent normalized confidence scores from visual, audio, and behavioral modules for event i . A weighted fusion score is computed as:

$$Fi = wvVi + waAi + wbBi \text{ where}$$

$wv + wa + wb = 1.$

Weights are selected empirically based on validation performance to balance sensitivity and robustness.

$$EIS = i = 1 \sum \_ nwi \cdot Ri$$

3.3.2 Liveness and Anti-Spoofing

To preventimpersonation attacks, liveness detection analyses: **a)** Eye blink frequency

b) Subtle facial muscle movements

c) Texture inconsistencies between real and spoofed faces

d) Depth cues when available

Only candidates passing both identity and liveness checks are permitted to begin the examination.

3.4 Computer Vision-Based Monitoring

3.7 Exam Integrity Scoring and Alerts

a) Candidate speech

b) Background conversations

c) Device alerts

d) Benign ambient noise Suspicious acoustic events are logged with corresponding confidence scores.

3.6 Behavioural Modelling and Multimodal Fusion

3.6.1 Temporal Behaviour Analysis

To capture behavioural trends over time, sequential models such as Long Short-Term Memory (LSTM) networks analyse patterns including:

a) Frequent gaze shifts

b) Repeated posture changes

c) Continuous movement anomalies

where is the number of detected events. The score is set from 0 which indicates high risk to 1 indicating low risk.

Alerts are classified as:

1) risk: Minor distractions
2) Medium risk: Face absence or identity mismatch
3) High risk: External assistance Low or device usage

3.8 Security, Privacy, and Ethical Compliance

All audio and video data are encrypted using AES-256 during transmission and storage. The system adheres to:

a) Minimal data retention policies

b) User consent mechanisms

c) Bias mitigation through diverse training data

d) Transparent reporting practices

Only authorized evaluators are permitted to access encrypted logs.

3.9 Testing, Validation, and Optimization System performance is evaluated using: **a)**Accuracy

b) Precision and recall

c) False positive rate

d) End-to-end response latency Pilot examinations are conducted under varying conditions to assess robustness. Pruning, quantization, and other model optimization techniques are applied to guarantee real-time performance on ordinary consumer devices.

3.10 Deployment Strategy

The framework supports deployment as:

a) A standalone desktop application

b) A browser-based platform

c) A hybrid client-server solution After each examination, automated integrity reports summarizing detected events and scores are generated for evaluators.

4. IMPLEMENTATION

This section explains how the proposed AI Powered Remote Assessment Monitor is practically developed and deployed. It describes how the architectural components discussed in the methodology are converted into a working system that can monitor examinations in real time. The main goal of the implementation is to create a solution that is reliable, scalable, and secure while remaining easy to use in real-world online examination environments. **4.1 System Realization and Objectives**

The implementation focuses on transforming the conceptual design of the system into a fully functional, software-based proctoring platform. The aim is to enable automated and continuous monitoring of candidates without requiring constant human supervision.

The system performs several key tasks during an examination. It verifies the identity of the candidate before the test begins, continuously observes visual and audio signals throughout the session, and analyzes these inputs to detect unusual or suspicious activities. Based on this analysis, it generates clear and interpretable integrity reports that help evaluators make informed decisions.

To ensure that the system can be used for large numbers of students, special attention is given to

modular design, low processing delay, and compatibility with regular consumer devices such as laptops or desktops. Lightweight models and optimized pipelines are therefore selected so that the system runs smoothly even on hardware with limited resources. A hybrid execution strategy is also adopted to balance local and server-side computation.

4.2 Modular System Architecture and Integration

The system is built using a modular architecture, where each component is responsible for a specific function but works together with other modules through a shared fusion engine. This separation of tasks will make it easier to maintain, upgrade, and scale the system in the future.

The main implementation modules are:

a) **Authentication Module** – Verifies the candidate's identity before the exam using facial recognition and liveness detection techniques.

b) **Vision Processing Engine** – Continuously analyzes the webcam feed

c) to detect face presence, eye gaze direction, head movement, and any unauthorized persons or objects.

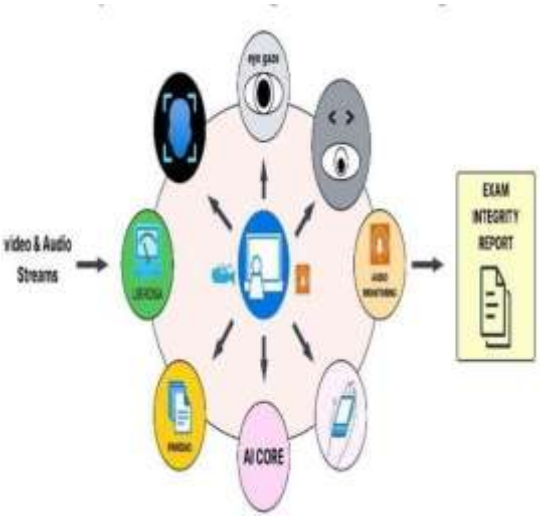
d) **Audio Analysis Engine** – Monitors background sounds and identifies suspicious noises or conversations during the examination.

e) **Behavioural Analysis Module** – Studies patterns over time using sequential models to better understand candidate behavior and reduce false alarms.

f) **Fusion and Scoring Engine** – Combines outputs from all modules and calculates an overall integrity risk score.

g) **Security Layer** – Prevents sensitive information by encrypting it and limiting access to authorized users only.

This modular structure ensures maintainability and enables parallel execution of different components. Also, maintaining or replacing individual modules without impacting the whole system is possible, thereby enhancing flexibility and scaling in the long-term.



4.3 Visual and Acoustic Processing Pipeline

Visual information collected from the candidate's webcam is processed continuously during the examination. From a video frame, essential features are identified, including the identification of the face, landmarks on the face, detection of eye rays, detection of head position, and the environment around the video frame. A lightweight object detection model is also used to identify unauthorized materials or the presence of additional people in the frame. These observations are then converted into normalized confidence scores that indicate the likelihood of potential integrity risks.

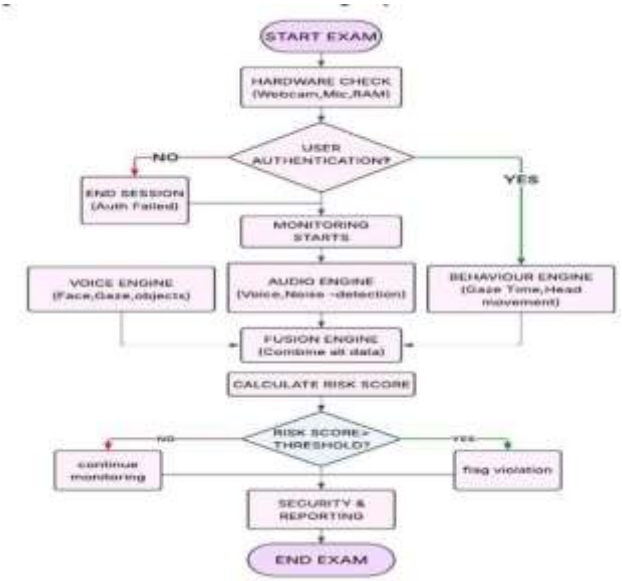


FIG. 3: System architecture flowchart showing the AI-

based pipeline that monitors video, audio, and behavioral signals to detect possible violations.

At the same time, the audio stream is analyzed in parallel. The sound is divided into short time segments and converted into frequency-based representations so that meaningful patterns can be captured. A convolutional neural network classifies each segment into categories such as candidate speech, background conversations, device sounds, or normal environmental noise. Only events that exceed predefined confidence levels are passed to the next stage, which helps minimize unnecessary alerts and reduces false positives.

4.4 Behavioural Modelling and Real-Time Fusion

Instead of judging candidate actions based on single, isolated events, the system focuses on understanding behavior over time. To achieve this, sequential neural networks are used to analyze patterns across multiple frames and audio segments. This helps the system identify repeated or sustained irregular activities that may suggest intentional misconduct rather than accidental movements.

The fusion engine then combines confidence scores from visual, audio, and behavioral modules using a weighted strategy. This ensures that no single source dominates the final decision. As a result, the system can ignore minor disturbances—such as brief glances away from the screen or occasional background noise—while still detecting consistent suspicious behavior more accurately.

4.5 Performance Optimization and Real-Time Constraints

To make sure the system runs smoothly on ordinary student devices, several optimization techniques are applied during implementation. These include model compression, limiting the frame rate, and batching audio data and video data to be sent for processing efficiently. Such optimizations do not drastically reduce detection accuracy while decreasing the amount of computation as much as possible.

The overall pipeline is constrained within realtime limits to ensure that monitoring is not only continuous but also responds to changes during the live exams. The performance trade- offs are relevant when deploying the system for mass use or on low hardware and limited bandwidth. **4.6 Security, Privacy, and Ethical Safeguards**

Since examination data is very sensitive; strong security measures are put in place in the entire system. All video, audio, and log data are encrypted both in transmission and storage to avoid unauthorized access. Also, only authorized personnel are are only is retained for the minimum period. Ethical considerations are also put into attention. In this system, there is clear consent from users; procedures, open monitoring policies, and Minimization of bias in the training of models. These steps help to insure fairnessness and trust Institutions.

4.7 Implementation Summary

AI Powered Remote Assessment (AIRAS) Monitor Proposed successfully converts the suggested design into a working solution that can be deployed. Combining video, audio, and behavioral analysis in a secure and optimized framework, the system proves that and have the reliability (accuracy) and validity (authentication) of a well-planned proctoring system that uses human proctors. Implementation emphasizes that scalable, real-time, and ethical online monitoring at scale, real-time, and, ethically, online examination modern AI technologies.

5. RESULTS

Testing under varied environmental conditions demonstrates reliable detection of suspicious activities with reduced false positives. The
multimodal strategy performs better than
singlemodality systems and maintains stable response times during live examinations.

6. CONCLUSION

The final product of the AI Powered Remote Assessment Monitor can be considered a significant advance in automated online proctoring. To offer a much smarter and more context-aware monitoring system, the AI Powered Remote Assessment Monitor does not rely on isolated rule-based checks but rather on a combination of multiple monitoring approaches. That ensures the provision of a level playing field, objectivity, and scalability of students in remote examinations.

A key strength of this work is the incorporation of video and audio analysis, which are integrated through a weighted fusion mechanism. Instead of

corresponding to basic yes-or-no alerts, a

weighted fusion mechanism produces an overall Exam Integrity Score (EIS), the "more nuanced and meaningful interpretation of candidate conduct" (ibid). Thus combined with continuously identity verification and anti spoofing, the exams set up is secure and limits the risks of impersonation or help outside the candidate.

In addition, the implementation demonstrates that successful AI-based monitoring need not violate student privacy or require costly hardware. Optimized processing and use of lightweight models mean the AI-backed monitoring model does not require expensive hardware to work efficiently; it can comfortably run on regular devices that are common among students. In combination, the AI Powered Remote Assessment Monitor provides a balance between improved security and practicality, accessibility, and fairness, thereby improving trust levels in the future of digital education.

7. FUTURE ENHANCEMENT

There are a number of opportunities to further improve the performance, fairness, and user acceptance of the AI Powered Remote Assessment Monitor, although the framework presents a strong and reliable multimodal online proctoring system already. The future modifications will focus not just on amplifying the level of technical accuracy, but also on making the system more transparent, privacyfriendlier, and more inclusive for all types of students.

7.1 Affective Computing and Stress Detection

Future versions are planned to include affective computing techniques to ascertain the emotional state of a candidate during examinations. Due to examination pressure, some students may be stressed or anxious, and this may be confused with suspicious behavior (ibid). A candidate's tiny micro-expressions and physiological cues like heart rate changes measured through remote photoplethysmography (rPPG) can be used to separate stress from deliberate cheating. This would help change the monitoring sensitivity and help eliminate the possibility of honest but nervous students being unfairly ranked as cheating the exam.

7.2 Explainable AI (XAI) for Transparency

Another important improvement should be making system's decisions clearer and more understandable. Instead of just giving the label "suspicious action", this information will be explained in a more comprehensible mannerusing visual indicators or brief comments. For example, suspiciousness could be indicated by the length of time during which the testee did not look at the monitor or background motion in the environment. This will allow the evaluators to better understand the alerts and acquire a higher degree of trust in the system's decisions.

7.3 Privacy-Preserving Federated Learning

Protecting student data is and remains its top priority. Protecting student data remains a key priority; hence, the system aims to adopt a federated learning approach where models are locally trained on individual devices instead of sending raw videos or audio data to a central server. Encrypted model deltas will only be shared. That way, biometric data stays strictly on the user's device, yet the system can still learn together across many institutions.

7.4 Accessibility and Threshold Customization

Future work will also include providing accessibility for students with some forms of physical or visual impediment. The standard thresholds used for monitoring may not apply to all people, especially those who have involuntary movements or users whose gaze is not usually at the center of the screen. Adaptive settings and customizable thresholds can be introduced to better tailor

the system to such needs and not penalize candidates for exhibiting them. This will help to ensure that the proctoring process remains inclusive and equitable for all users.

REFERENCES

[1] S. Saha, S. Sridevi, and J. C. K. Mani, "An AIbased intelligent exam proctoring system for secure and fair online assessments," J. Adv. Res. Arif. Intell. Appl., vol. 2, no. 1, pp. 1-7, Dec. 2024. [Online]. Available:https://www.researchgate.net/publication/386430495_An_AIBased_Intelligent_Exam_Proctoring_System_for_Secure_and_Fair_Online_Assessme_nts

[2] A. K. Naveen et al., "AutoOpen -- A multimodal framework for online exam proctoring," Sep. 2025, arXiv:2509.10887. [Online]. Available: https://arxiv.org/abs/2509.10887

[3] X. Li et al., "A visual analytics approach to facilitate the proctoring of online exams," Jan. 2021, arXiv:2101.07990. [Online]. Available: https://arxiv.org/abs/2101.07990

[4] G. Acapnia, "Detecting AI-assisted cheating in online exams through behavioral analytics," Sep. 2024, arXiv:2409.16923. [Online]. Available: https://arxiv.org/abs/2409.16923

[5] R. Wankhade et al., "Temporal analysis for automated proctoring systems," Oct. 2025, arXiv:2510.18881. [Online]. Available: https://arxiv.org/abs/2510.18881

[6] Y. Chen et al., "Visual analytics for behavior monitoring in remote examinations," Jun. 2022, arXiv:2206.13356. [Online]. Available: https://arxiv.org/abs/2206.13356

[7] A. Sridhar and J. S. Rajshekhar, "AIintegrated proctoring system for online exams," J. Arif. Intell. Capsule Newt., vol. 4, no. 2, pp. 139-148, 2022. [Online]. Available: https://irojournals.com/aicn/article/view/4/2/6

[8] E. Xu, J. Lu, S. Xu, and J. Wang, "Cheating recognition in examination halls based on improved YOLOv8," Discover Computer, vol. 28, no. 1, Art. no. 256, Dec. 2025. [Online]. Available: https://link.springer.com/article/10.1007/s107910 25-09747-3

[9] J. R. Pansare, A. Pawar, A. Chorghade, S. Barge, and A. Agarwal, "Proctoring using AI," Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET), vol. 13,no. 6, Jun. 2025. [Online]. Available:https://www.ijraset.com/researchpape r/ proctoring-using-ai

[10] Niharika G. N. and S. N. Nayak, "Artificial intelligence based online examination proctoring system," Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET), vol. 11, no. 9, pp. 569-573, Sep. 2023. [Online]. Available:https://www.ijraset.com/researchpaper/ artificial-intelligence-based-onlineexaminationproctoring-system