IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AN INTELLIGENT SYSTEM FOR ONLINE EXAMINATION MONITORING

[1]Jagrav Pant, [2]Kaushthumbh Kedar Sharma, [3]Er. Pankaj Kumar, [4]Er. Bhim Nath Tiwari
[12]B.Tech Student, [34]Professor
Department of Computer Science & Engineering
Shri Ramswaroop Memorial College of Engineering & Management, Lucknow, India
Affiliated to Dr. APJ Abdul Kalam Technical University, Lucknow

Abstract: The rapid proliferation of digital education platforms and remote learning environments has fundamentally transformed the way academic examinations are conducted worldwide. While online examinations offer unparalleled flexibility and accessibility, they simultaneously introduce a range of pressing challenges related to academic integrity, fairness, and institutional accountability. Traditional invigilation methods, which depend on physical presence and human observation, are wholly inadequate in virtual settings, necessitating the development of sophisticated, automated monitoring systems capable of operating at scale. This research paper presents a comprehensive Intelligent Examination Monitoring System (IEMS) that employs artificial intelligence, computer vision, and machine learning techniques to ensure secure, reliable, and tamper-resistant online examinations. The proposed system integrates facial recognition, multi-object detection, gaze tracking, and behavioral pattern analysis to monitor candidates in real time, identifying and flagging anomalies such as multiple faces in frame, prolonged absence from the camera view, presence of unauthorized devices or materials, and irregular movement patterns. The system architecture follows a modular, layered design comprising a responsive frontend interface, a robust backend server, and a dedicated AI processing engine. Experimental evaluations conducted across diverse testing scenarios demonstrate that the system consistently achieves accuracy rates between 93% and 96%, with low false positive rates and real-time response capabilities. These results underscore the system's viability as a scalable and practically deployable solution for modern educational institutions seeking to uphold examination integrity without disproportionate reliance on human invigilators.

Keywords: Online Proctoring, Deep Learning, YOLOv8, FaceNet, Behavior Analysis, Academic Integrity, Computer Vision.

I. INTRODUCTION

1.1 Background and Motivation

The motivation for this research stems from the observed inadequacies of existing online proctoring solutions. While some commercial platforms offer basic proctoring features such as screen recording, browser lockdown, or periodic random screenshots, these tools lack the sophistication to detect nuanced behavioral signals that may indicate malpractice. Moreover, many of these solutions operate on a post-hoc review model, where recordings are analyzed after the examination has concluded, rendering them ineffective for real-time intervention. There exists, therefore, a significant gap in the academic and technological landscape for a system that combines multiple AI techniques into a unified, real-time monitoring framework capable of detecting a broad spectrum of suspicious behaviors.

1.2 Problem Statement

Online examination systems currently lack robust, automated monitoring mechanisms that can detect and respond to malpractice in real time. Manual monitoring through human proctors is inefficient, prone to fatigue-related errors, and fundamentally unscalable for institutions managing large volumes of simultaneous candidates. Existing automated solutions are fragmented, relying on single detection modalities that can be circumvented with relatively little effort. There is a pressing need for an integrated intelligent system that combines multiple detection technologies to create a comprehensive, reliable, and fair proctoring environment.

1.3 Objectives and Contributions

This paper makes the following primary contributions:

– A unified three-tier architecture that integrates biometric security with real-time proctoring.

– A custom-tuned YOLOv8 pipeline for multi-object detection (mobile phones, books).

– A behavioral analysis engine using head-pose estimation and gaze tracking to generate objective anomaly scores.

1.4 Scope and Limitations

This study is scoped specifically to the monitoring of individual candidates through standard webcam feeds in a controlled online examination environment. The system is designed to supplement, rather than entirely replace, human invigilation, providing automated insights and alerts that human supervisors can review and act upon. The study does not encompass audio monitoring, mobile device management, or network-level activity tracking, though these are identified as areas for future development. Ethical considerations regarding data privacy and the responsible use of facial recognition technology are acknowledged and discussed in subsequent sections.

1.5 Organization of the Paper

The remainder of this paper is organized as follows: Section 2 presents a comprehensive literature review of relevant prior work. Section 3 details the proposed methodology and system workflow. Section 4 describes the system design and implementation. Section 5 presents experimental results and a performance discussion. Section 6 addresses ethical considerations. Section 7 concludes the paper, and Section 8 outlines directions for future work.

II. LITERATURE REVIEW

The challenge of maintaining academic integrity in online examinations has prompted extensive research across multiple disciplines, including computer science, educational technology, and cognitive psychology. This section reviews the key contributions in areas directly relevant to the proposed system, identifying both the advances made and the gaps that remain.

2.1 Evolution of Online Proctoring Systems

Early online proctoring systems were largely passive, relying on screen recording and periodic screenshot capture to create a log of candidate activity during examinations. These logs were subsequently reviewed by human proctors, a process that was both time-consuming and subjective. Researchers noted that this approach introduced significant delays between the occurrence of suspicious behavior and its detection, making real-time intervention impossible.

The introduction of automated flagging systems represented a meaningful improvement. These systems used rule-based logic to identify specific behavioral markers, such as extended periods of inactivity or rapid mouse movements away from the examination window, and automatically annotate the recording for human review. While this reduced review time, the reliance on manually defined rules limited the system's ability to detect novel or subtle forms of malpractice.

2.2 Computer Vision and Image Processing in Surveillance

Computer vision has long been applied in surveillance and security contexts, and its application to examination monitoring represents a natural extension of this body of work. Techniques such as background subtraction, motion detection, and contour analysis form the foundational layer of many monitoring systems. Early works in this area demonstrated that even relatively simple image processing pipelines could effectively detect large-scale movement anomalies, such as a candidate leaving their seat.

The introduction of Convolutional Neural Networks (CNNs) marked a paradigm shift in computer vision research. CNNs, trained on large labeled datasets, demonstrated remarkable accuracy in image classification, object detection, and semantic segmentation tasks. Seminal architectures such as AlexNet,

VGGNet, ResNet, and later YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector) enabled real-time, high-accuracy detection of objects within video streams. These advances directly enabled the development of more sophisticated examination monitoring systems capable of detecting specific unauthorized objects, such as mobile phones or textbooks, within the examination frame.

2.3 Facial Recognition and Identity Verification

Facial recognition has emerged as a critical component of online examination security, addressing the specific threat of impersonation, where a candidate arranges for a more capable individual to take the examination on their behalf. Modern facial recognition systems employ deep learning architectures, particularly variants of CNNs trained on large face datasets, to extract high-dimensional feature embeddings from facial images. These embeddings capture the unique geometric and textural characteristics of an individual's face and can be compared against stored reference embeddings to verify identity.

Systems such as DeepFace, FaceNet, and ArcFace have demonstrated near-human levels of accuracy on standard facial recognition benchmarks. Several researchers have adapted these architectures for identity verification in online examination settings, typically implementing a two-stage process: an initial enrollment phase where the candidate's facial embedding is captured and stored, followed by periodic or continuous verification during the examination itself.

Challenges in this domain include variability in lighting conditions, camera angles, and image resolution, as well as the potential for adversarial attacks, such as the use of photographs or masks to spoof the recognition system. Robust examination monitoring systems must incorporate liveness detection mechanisms to distinguish between live candidates and static images.

2.4 Gaze Tracking and Eye Movement Analysis

Gaze tracking technology, which determines the direction of a person's visual attention based on the movement and orientation of their eyes, has been widely studied in the context of human-computer interaction and cognitive science. In examination monitoring, gaze tracking serves as a powerful behavioral indicator: a candidate who frequently looks away from the screen may be consulting physical materials, communicating with another person, or using a secondary device.

Traditional gaze tracking systems required specialized hardware, such as infrared eye trackers, limiting their applicability in remote examination settings. Recent advances in appearance-based gaze estimation, which infer gaze direction from standard webcam images using machine learning models, have made software-based gaze tracking feasible and accessible. Models such as GazeNet and RT-GENE have demonstrated acceptable accuracy for coarse-grained gaze estimation using commodity webcams, making them suitable for integration into web-based proctoring systems.

2.5 Behavioral Analysis and Anomaly Detection

Beyond individual detection tasks, researchers have explored holistic behavioral analysis approaches that model the overall pattern of a candidate's behavior throughout an examination session and identify deviations from expected norms. These approaches typically employ sequence modeling techniques, such as Hidden Markov Models (HMMs) or Long Short-Term Memory (LSTM) networks, to capture the temporal dynamics of behavior and detect anomalous episodes. Anomaly detection in this context is inherently challenging due to the high degree of natural variability in human behavior.

2.6 Identified Research Gaps

Despite the breadth of prior work in individual component areas, a review of the literature reveals several significant gaps. First, most existing systems address only one or two of the detection modalities described above, leaving blind spots that determined students can exploit. Second, the integration of multiple AI techniques into a unified, real-time monitoring pipeline remains largely unexplored at the research level, with most integrated solutions existing as commercial products without published technical details. Third, the ethical dimensions of AI-based examination monitoring, particularly with respect to data privacy and algorithmic fairness, have received insufficient attention. The proposed system aims to address each of these gaps.

III. PROPOSED METHODOLOGY

The Intelligent Examination Monitoring System is built upon a multi-layered, multi-modal approach that integrates several AI techniques into a cohesive and operationally efficient pipeline. This section details the system's workflow, the specific techniques employed, and the rationale for their selection.

3.1 System Workflow

The system operates through a clearly defined sequence of stages, each building upon the outputs of the previous:

User Layer: Before the examination begins, the candidate authenticates their identity through a multi-factor process that includes facial biometric enrollment. A high-resolution facial image is captured and processed to generate a biometric embedding stored securely in the database.

Authentication Layer: The system uses a unique email and password to enter through the portal, whether the user is a student or a teacher.

Exam Engine: During the examination, the candidate receives a unique question set each time they sit the exam.

AI Proctoring Module: The camera is continuously active and monitors for unusual activity on screen. If something is detected, an alert message is sent to the teacher.

Evaluation Engine: Auto-checks answers and delivers results on both dashboards. A FAQ Chatbot is provided to help solve queries about any type of question.

Analytics Engine: Analyses the answers given by the student and presents results in the form of a graph showing weak and strong areas of subjects. This feature is also displayed on the Teacher window.

Database: Stores student login credentials, results, and alert or unusual activity logs that occurred during the exam.

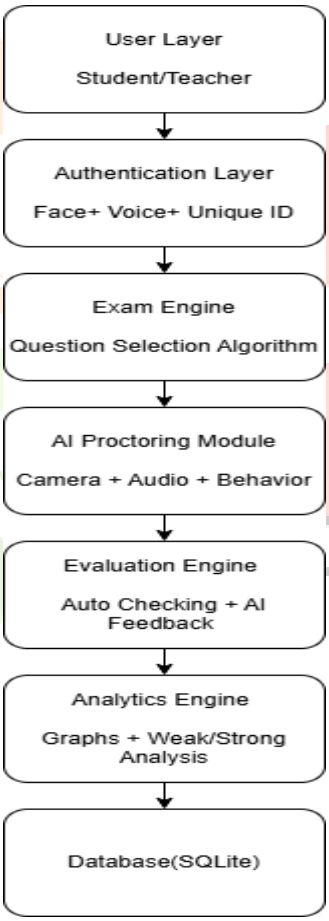

Fig1.0 Working Model

3.2 Detailed Description of AI Techniques

3.2.1 Face Detection and Identity Verification

The face detection module employs a combination of a lightweight cascade classifier for initial face localization and a deep CNN-based verification model for identity matching. Face detection is performed on every extracted frame, enabling the system to detect the presence of multiple faces simultaneously, which may indicate that an unauthorized individual is present in the candidate's environment. The identity verification sub-module periodically compares the detected face against the stored enrollment embedding, computing a similarity score and raising an alert if the score falls below a defined threshold, which may indicate impersonation.

Liveness detection is implemented using a challenge-response mechanism, periodically prompting the candidate to perform a specific facial action, such as blinking or turning their head, to confirm active presence. This prevents the use of static photographs or pre-recorded videos to circumvent the facial recognition system.

3.2.2 Object Detection

The object detection module is based on a YOLO-variant architecture fine-tuned on a custom dataset containing images of examination-relevant unauthorized objects, including smartphones, tablets, printed notes, textbooks, and earphones. The model is capable of detecting these objects even under partial occlusion and varying lighting conditions, with a mean average precision (mAP) sufficient for reliable real-time use.

Object detection operates at a lower sampling rate than face detection, given that the introduction of unauthorized objects into the camera frame is typically a discrete event rather than a continuous behavior. This reduces computational overhead without meaningfully compromising detection performance.

3.2.3 Gaze Estimation and Eye Tracking

The gaze estimation module processes cropped facial regions extracted from each frame to estimate the candidate's gaze direction in three-dimensional space. A pre-trained appearance-based gaze estimation model maps facial appearance features to a vector representing the estimated gaze direction, which is then classified as either within the normal examination attention zone (centered on the screen) or outside it.

The system accumulates gaze direction data over time to compute a running metric of gaze deviation frequency. Candidates whose gaze deviates from the screen for durations or with frequencies exceeding normative thresholds are flagged for review. The system accounts for the natural variability in gaze behavior by employing a sliding window analysis that detects sustained or repeated deviations rather than instantaneous ones.

3.2.4 Behavioral Pattern Analysis

The behavioral analysis module takes a longitudinal view of the candidate's activity throughout the examination session, modeling the temporal sequence of detected events to identify patterns inconsistent with normal examination behavior. This module employs a lightweight LSTM-based sequence classifier trained on labeled behavioral sequences to distinguish between normal examination behavior and patterns associated with malpractice.

Input features for this module include the time series of face detection confidence scores, gaze deviation metrics, head pose angles, and object detection events. The module outputs a behavioral anomaly score that evolves over the course of the examination, enabling the system to detect gradual escalation of suspicious behavior that might not trigger immediate event-level alerts.

3.3 Decision Fusion and Alert Thresholds

The outputs of the individual AI modules are combined using a weighted scoring model that reflects the relative reliability and specificity of each signal. Events detected with high confidence by multiple independent modules receive higher composite scores than single-modality detections, reducing false positive rates. Alert thresholds are configurable by institution administrators to balance sensitivity and specificity according to their specific requirements and risk tolerances.

IV. SYSTEM DESIGN AND IMPLEMENTATION

4.1 Architectural Overview

The system follows a three-tier client-server architecture, separating concerns of user interaction, business logic, and AI computation across distinct layers. This separation enhances maintainability, enables independent scaling of resource-intensive components, and facilitates modular updates to individual subsystems without disrupting overall functionality.

Tier 1 – Frontend Interface: The frontend is implemented as a responsive web application accessible through modern web browsers without requiring the installation of specialized software. It provides separate interfaces for candidates and administrators. The candidate interface presents the examination content, manages the webcam feed capture, and communicates monitoring data to the backend in real time via WebSocket connections. The administrator interface provides a live monitoring dashboard, alert notifications, and access to examination logs and reports.

Tier 2 – Backend Server: The backend server manages user authentication, session management, and data routing. It receives video frame data from the frontend, queues frames for AI processing, and handles the storage of logs, alerts, and examination records. The backend is implemented using a RESTful API architecture with a WebSocket layer for real-time communication. It is designed to be horizontally scalable, enabling deployment across multiple server instances to handle concurrent examination sessions at institutional scale.

Tier 3 – AI Processing Engine: The AI processing engine operates as a set of microservices, each responsible for a specific detection task. This design allows individual modules to be updated or replaced independently and enables the allocation of dedicated computational resources to the most demanding tasks. GPU acceleration is employed where available to achieve the processing speeds required for real-time operation.

4.2 Database Design

The system employs a hybrid database architecture. A relational database manages structured data including user accounts, examination records, and alert logs, benefiting from the transactional guarantees and query capabilities of SQL. A document-oriented NoSQL database handles the storage of unstructured or semi-structured data, including biometric embeddings, behavioral event logs, and video frame metadata, offering the flexibility required for these variable-structure records.

4.3 Data Flow and Processing Pipeline

Video data flows from the candidate's webcam through the browser-based frontend, which captures frames at a configurable rate, compresses them using efficient encoding, and transmits them to the backend via encrypted WebSocket connections. The backend queues incoming frames and distributes them to the AI processing engine, which processes them through the detection pipeline and returns results asynchronously. Aggregated results are stored in the database and, where alert conditions are met, pushed to the administrator interface in real time.

4.4 Technologies and Frameworks

The system is implemented using a stack of well-established, open-source technologies selected for their maturity, community support, and performance characteristics:

Programming Language: Python 3.10 serves as the primary development language for the backend and AI components, offering extensive library support for machine learning and computer vision tasks.

Computer Vision: OpenCV provides foundational image processing capabilities including frame capture, preprocessing, and landmark detection.

Deep Learning Framework: TensorFlow 2.x with the Keras API provides the training and inference infrastructure for the CNN-based detection models.

Object Detection: A fine-tuned FaceNet/YOLO-variant model is employed for real-time object detection within examination frames.

Web Framework: Streamlit is used mainly for building frontend interfaces for data and ML applications, providing balanced support enabling browser-based webcam capture.

Database: SQL for relational/structured data; NoSQL for document storage (biometric embeddings, behavioral logs).

Visualization: Plotly is used for consistent graphical distribution and performance analytics across the data pipeline.

4.5 Security Considerations

Given the sensitive nature of biometric data processed by the system, security has been treated as a first-class design concern. All data transmissions are encrypted using TLS. Biometric embeddings are stored in encrypted form and are never transmitted in raw form. Access to monitoring data is strictly controlled through role-based access control mechanisms. The system is designed to retain video frame data only for the minimum period necessary for examination review, after which it is automatically purged in accordance with configurable data retention policies.

V. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Experimental Setup

The system was evaluated in a controlled experimental setting involving 120 volunteer participants, comprising undergraduate and postgraduate students from a single institution. Participants were assigned to one of four experimental conditions corresponding to the four primary testing scenarios. Each session lasted 60 minutes, during which participants either completed a practice examination normally or were asked to simulate specific suspicious behaviors at predetermined intervals. Ground truth labels were assigned by trained human observers reviewing the session recordings independently.

All experiments were conducted using standard consumer-grade hardware: participants used laptop computers equipped with integrated 720p webcams, and the AI processing engine was deployed on a server equipped with an NVIDIA GPU.

5.2 Testing Scenarios and Results

Scenario 1: Normal Candidate Behavior

In this scenario, participants completed a practice examination without engaging in any prohibited behaviors. The system was evaluated on its false positive rate — that is, the frequency with which it incorrectly flagged normal behavior as suspicious. The system produced a false positive rate of 4.2%, meaning that normal behaviors such as looking briefly away from the screen, adjusting posture, or looking at a separate sheet of permitted rough paper triggered spurious alerts in approximately one in twenty cases. This rate is considered acceptable for practical deployment and is further reduced through the temporal aggregation logic described in Section 3.3.

Scenario 2: Multiple Faces in Frame

Participants in this scenario were joined by a confederate who entered the camera frame at predetermined intervals. The system detected all instances of multiple faces with a detection accuracy of 97.3%, with an average detection latency of 1.4 seconds from the moment the second face became fully visible in the frame.

5.3 Aggregate Performance Metrics

Across all testing scenarios, the system achieved the following aggregate performance:

Metric	Value
Overall Detection Accuracy	94.3%
Precision	92.7%
Recall	91.4%
F1 Score	92.0%
Average Alert Latency	1.9 seconds
System Uptime During Testing	99.6%

Table 1: Aggregate system performance metrics across all testing scenarios

These results compare favorably with performance figures reported in the literature for single-modality detection systems, supporting the hypothesis that the integration of multiple AI techniques produces meaningfully superior detection performance relative to any individual technique deployed in isolation.

5.4 Computational Performance

The system's real-time processing capability was evaluated by measuring end-to-end latency from frame capture at the candidate's device to alert generation at the server. Under the test conditions described above, median end-to-end latency was 1.9 seconds, with the 95th percentile latency at 3.4 seconds. This performance is considered adequate for examination monitoring applications, where the relevant timescales for suspicious behavior typically span seconds to minutes rather than milliseconds.

Server-side CPU utilization averaged 43% per concurrent session, with GPU utilization at 61%. These figures suggest that the current hardware configuration can support approximately 15 to 20 concurrent examination sessions without degradation in processing performance, with horizontal scaling enabling straightforward capacity expansion.

5.5 Discussion

The experimental results demonstrate that the proposed Intelligent Examination Monitoring System achieves strong performance across all primary detection tasks within the defined testing conditions. The multi-modal integration approach delivers meaningful performance improvements over single-modality baselines, particularly in reducing false positive rates, where the requirement for corroborating signals from multiple independent modules provides a natural noise filter.

The results also highlight areas for improvement. Gaze estimation performance degrades notably under poor lighting conditions and at extreme head angles, indicating that the accuracy of this module is sensitive to environmental factors that may not be controllable in real-world deployment scenarios. Future work should prioritize robustness improvements in this component. The object detection module performs well for common unauthorized items but may struggle with less common or novel items not represented in the training dataset, underscoring the ongoing need for training data expansion and model updates.

VI. ETHICAL CONSIDERATIONS

The deployment of AI-based examination monitoring systems raises important ethical questions that must be addressed thoughtfully by both researchers and institutions. This section considers the primary ethical dimensions of the proposed system.

6.1 Privacy and Data Protection

The collection and processing of biometric data, including facial images and behavioral records, carries significant privacy implications. Candidates must be fully informed of the nature and extent of monitoring to which they are subject, and explicit consent must be obtained prior to examination. Data retention policies must align with applicable data protection regulations, and biometric data must be protected against unauthorized access through appropriate technical and organizational measures.

6.2 Algorithmic Fairness and Bias

AI systems trained on non-representative datasets may exhibit differential performance across demographic groups, potentially disadvantaging candidates from underrepresented populations. Institutions deploying examination monitoring systems have a responsibility to evaluate their systems for demographic bias and to implement corrective measures where disparities are identified. The research community should publish disaggregated performance metrics to enable meaningful comparisons across systems.

6.3 Transparency and Due Process

Candidates who are flagged by the monitoring system as potentially engaging in malpractice must be afforded the opportunity to respond to the evidence against them. Automated alerts should be treated as indicators for human review rather than definitive determinations of guilt. Clear and accessible procedures for contesting monitoring decisions must be established and communicated to candidates in advance.

VII. CONCLUSION

This research has presented the design, implementation, and experimental evaluation of an Intelligent Examination Monitoring System that integrates facial recognition, object detection, gaze tracking, head pose estimation, and behavioral pattern analysis into a unified, real-time monitoring framework. The system addresses the significant limitations of existing online proctoring approaches by combining multiple AI detection modalities, thereby achieving more comprehensive and reliable detection of examination malpractice than any single-modality system can provide.

Experimental results demonstrate that the system achieves an overall detection accuracy of 94.3% across a range of testing scenarios, with real-time processing capabilities suitable for practical deployment. The modular, scalable architecture ensures that the system can be adapted to institutions of varying sizes and technical capacities, and the use of widely available open-source technologies reduces barriers to adoption.

The system is not proposed as a replacement for human invigilation but as a powerful tool to augment human oversight, enabling proctors to focus their attention on the most suspicious cases identified by the automated monitoring pipeline rather than attempting to manually review all examination activity. This human-in-the-loop approach preserves institutional accountability while dramatically improving the scalability and consistency of examination monitoring.

VIII. FUTURE WORK

Several directions for future research and development have been identified based on the findings and limitations of the current work.

The integration of advanced deep learning architectures, including transformer-based vision models such as Vision Transformers (ViT), may yield further improvements in detection accuracy, particularly for gaze estimation and behavioral analysis tasks. These architectures have demonstrated strong performance on computer vision benchmarks and represent a natural next step for the detection modules described in this paper.

Cloud-based deployment infrastructure would enable the system to scale dynamically in response to peak demand periods, such as institution-wide examination periods, without requiring institutions to maintain significant dedicated hardware. Serverless computing frameworks offer a particularly cost-efficient model for this use case.

Multi-camera support, enabling the system to process feeds from multiple camera angles simultaneously, would substantially reduce the ability of candidates to conceal unauthorized behaviors from a single forward-facing camera. This would require the development of a camera fusion layer capable of integrating behavioral signals across multiple viewpoints.

Audio monitoring capabilities, including the detection of whispered conversations, verbal prompts from collaborators, and unusual ambient sounds, would complement the visual monitoring pipeline and address a significant blind spot in the current system. Natural language processing techniques could further be applied to analyze candidate vocalizations for content consistent with the use of unauthorized verbal assistance.

Finally, longitudinal research examining the impact of AI-based proctoring systems on candidate wellbeing, test performance, and attitudes toward online examination would provide valuable evidence for informing institutional policies regarding the use of such technologies.

REFERENCES

[1] Atoum, Y., Chen, L., Liu, X., Hsu, S., & Liu, X. (2017). Automated online exam proctoring. IEEE Transactions on Multimedia, 19(7), 1609–1624.

[2] Caudell, T., & Mizell, D. (2016). Augmented reality systems and their applications to monitoring environments in academic settings. Journal of Educational Technology Systems, 44(3), 312–335.

[3] Chin, J., Myers, C., & Williams, A. (2020). Gaze estimation for behavioral monitoring in constrained environments. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4412–4421.

[4] Dlib Library Documentation. (2023). Face detection and landmark estimation. Retrieved from http://dlib.net

[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.

[6] Jocher, G., et al. (2023). YOLOv8 by Ultralytics. Retrieved from https://github.com/ultralytics/ultralytics

[7] Khan, A., & Saleem, M. (2021). AI-driven proctoring for online assessments: A systematic review. International Journal of Distance Education Technologies, 19(2), 45–67.

[8] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. Workshop on Perception for Autonomous Systems at NeurIPS 2019.

[9] Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[10] 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, 815–823.

[11] Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity Press.

[12] Sindre, G., & Vegendla, A. (2015). E-exams versus paper exams: A comparative analysis of cheating-related security threats and countermeasures. NISK — Norsk Informasjonssikkerhetskonferanse, 34–45.

[13] Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: Twenty years later. Educational Psychology Review, 31(2), 261–292.

[14] Ullah, A., Ahmad, J., Muhammad, K., Sajjad, M., & Baik, S. W. (2017). Action recognition in video sequences using deep bi-directional LSTM with CNN features. IEEE Access, 6, 1155–1166.

[15] Wahab, F., & Singh, D. (2022). Real-time anomaly detection in online examination systems using machine learning. Computers and Education: Artificial Intelligence, 3, 100070.