International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

AI-Based Online Proctoring System for Secure and Scalable Remote Examinations

Mayur Patil, Kunal Viroje, Harsh Waingankar, Dr. Vivek Khalane, Dr. Vaibhav Narawade
Department of Computer Engineering Ramrao Adik Institute of Technology
Navi Mumbai, India

Abstract- Online examinations have become a common part of modern education, especially with the growth of remote learning platforms. However, maintaining fairness and preventing malpractice in such environments remains a major challenge. In this work, we present an AI-based online proctoring system designed to monitor candidates during examinations using real-time video and audio analysis. The system combines face recognition, gaze tracking, head pose estimation, and audio monitoring to detect suspicious activities such as impersonation, presence of multiple individuals, and abnormal behavior. During our testing across multiple sessions and varying environmental conditions, we observed that the system achieved an overall detection accuracy of approximately 92.6% while maintaining real-time performance of 20–30 frames per second. The proposed system reduces dependency on human invigilators and provides a scalable solution for large-scale online examinations.

Keywords- Online Proctoring, Artificial Intelligence, Computer Vision, Face Recognition, Exam Integrity, Behavioral Analysis.

I. INTRODUCTION

The rapid advancement of digital technologies and widespread internet accessibility have significantly transformed the education sector. In recent years, online learning platforms have gained immense popularity, enabling students to access educational resources and assessments from any location. This shift became even more prominent during global disruptions such as the COVID-19 pandemic, where institutions were forced to adopt remote learning and evaluation methods. While online examinations offer flexibility, scalability, and convenience, they also introduce serious concerns related to academic integrity and fairness.

One of the primary challenges in online examinations is the increased possibility of malpractice. Unlike traditional classroom environments, remote exams lack physical supervision, making it easier for candidates to engage in activities such as impersonation, consulting unauthorized materials, or receiving external assistance. Conventional methods of online proctoring, such as live human invigilation through webcams, are not only resource- intensive but also difficult to scale for large numbers of candidates. Moreover, human monitoring is prone to fatigue, bias, and inconsistency, which can affect the reliability of the evaluation process.

To address these challenges, there has been growing interest in the application of Artificial Intelligence (AI) and Computer Vision techniques for automated proctoring. AI-based systems can continuously monitor candidate behavior, analyze patterns, and detect anomalies in real time without requiring constant human intervention. These systems utilize technologies such as facial recognition for identity verification, gaze tracking for attention monitoring, head pose estimation for movement analysis, and audio processing for detecting background conversations or suspicious sounds.

In our work, we focus on designing and implementing an AI-based online proctoring system that aims to provide a reliable, scalable, and efficient solution for remote examinations. During the development and testing phase, we observed that combining multiple behavioral indicators significantly improves the accuracy of anomaly detection compared to relying on a single method. For instance, integrating gaze tracking with head movement analysis helps in better identifying attention shifts, while combining audio and visual cues reduces false positives.

The proposed system captures real-time video and audio data from the candidate's device and processes it using a combination of machine learning models and computer vision algorithms. Unlike many existing solutions that rely heavily on cloud-based processing, our approach is designed to be efficient and adaptable to varying system capabilities. The system generates real-time alerts for suspicious activities and maintains a detailed log of events, which can be reviewed by examiners after the examination.

The main objective of this work is to enhance the credibility and fairness of online examinations while minimizing the need for human intervention. By providing automated, consistent,

and real-time monitoring, the system aims to support educational institutions in conducting secure and large-scale online assessments. Additionally, this research contributes to the ongoing development of intelligent surveillance systems in education, paving the way for future advancements in remote learning technologies.

II. PROBLEM STATEMENT

The increasing use of online examinations has introduced significant challenges in maintaining fairness and academic integrity. Unlike traditional exam environments, remote assessments lack direct physical supervision, making it easier for candidates to engage in malpractice such as impersonation, use of unauthorized materials, or receiving external assistance. Existing proctoring solutions attempt to address these issues through human invigilation or automated monitoring. However, human-based supervision is not scalable for large numbers of candidates and is often affected by fatigue, inconsistency, and subjective judgment. On the other hand, many automated systems rely on basic detection methods such as face presence or browser activity, which are insufficient to identify subtle behaviors like gaze diversion, minor head movements, or background communication.

Another major limitation is the lack of effective real- time detection. Several systems analyze recorded data only after the exam, which prevents immediate action against suspicious activities. Additionally, many solutions depend on continuous high-bandwidth internet connectivity, making them unreliable in environments with unstable network conditions.

Privacy concerns also play a crucial role, as continuous monitoring involves capturing sensitive audio and video data, raising issues related to data security and user trust.
Therefore, there is a need for a scalable, accurate, and real-time online proctoring system that can detect a wide range of suspicious behaviors using multiple techniques while minimizing human intervention and ensuring privacy and reliability.

III. LITERATURE SURVEY

In recent years, significant research has been conducted in the field of online examination proctoring, particularly with the integration of Artificial Intelligence and machine learning techniques. One of the foundational works in this domain was proposed by Atoum et al., who developed an automated online proctoring system using multimedia analytics. Their system utilized video and audio inputs to detect behaviors such as gaze movement, voice activity, and use of unauthorized devices, demonstrating strong accuracy in identifying cheating activities [1].

Further research has focused on improving identity verification and robustness under real-world conditions. Hadian and Bandung proposed a continuous user verification approach that enhances face recognition performance even under varying lighting and head pose conditions, making it more suitable for practical deployment [2]. In addition, modern systems have incorporated deep learning models such as convolutional neural networks and object detection frameworks like YOLO to identify multiple persons and suspicious objects in real time, significantly improving monitoring capabilities [3].

A comprehensive review by Nigam et al. analyzed multiple AI-based proctoring systems and emphasized the importance of multi-modal approaches. Their study highlighted that combining facial recognition, gaze tracking, and audio analysis leads to higher detection accuracy compared to single-modality systems. However, the review also pointed out challenges such as privacy concerns, ethical implications, and high computational requirements [4].

Recent advancements have explored integrated frameworks that combine multiple behavioral cues to generate real-time alerts and cheating probability scores. These systems show improved performance but often require more computational resources and optimized architectures to maintain real-time efficiency [5].

Overall, existing research demonstrates that while AI- based proctoring systems have made considerable progress, there are still gaps in terms of scalability, real- time performance, and privacy preservation, which need to be addressed for practical large-scale deployment.

IV. RESEARCH GAP

Despite significant advancements in AI-based proctoring systems, several limitations still exist. Many systems lack real-time detection capabilities, rely on single-modality analysis, and depend heavily on stable network conditions. Additionally, privacy concerns and high computational requirements limit their practical deployment.
This motivates the need for a scalable, multi-modal, and efficient system that can operate reliably in real-world environments.

V. SYSTEM DESIGN

The proposed online proctoring system is designed as a sequential processing pipeline that analyzes candidate behavior in real time. As illustrated in the architecture diagram, the system begins with capturing live video input from the user's webcam. The captured video is divided into individual frames

through a frame capture module, which enables continuous analysis.

These frames are then passed to the AI detection module, which acts as the central processing unit of the system. This module distributes the input to multiple detection components for parallel analysis. The face detection module verifies the presence and identity of the candidate, ensuring that no unauthorized person appears in the frame. Simultaneously, the gaze tracking module monitors eye movement to detect attention shifts or suspicious viewing directions. The audio analysis module processes sound input to identify background noise or conversation, which may indicate malpractice.

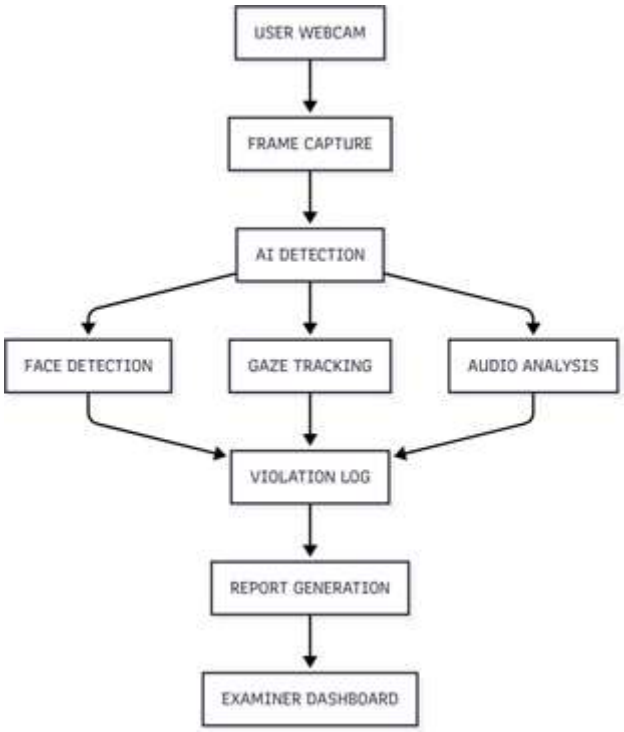

Fig.1. Architecture Flow

The outputs from all detection modules are combined and forwarded to the violation logging module, where suspicious activities are recorded along with timestamps. This ensures that all detected anomalies are systematically documented for later review.

Following this, a report generation module compiles the recorded violations into a structured summary, highlighting key events during the examination session. Finally, the processed information is displayed on the examiner dashboard, allowing invigilators to review candidate behavior efficiently.

This pipeline-based design ensures real-time performance, modularity, and scalability, making the system suitable for large-scale online examination environments.

VI. METHODOLOGY

The proposed system follows a multi-stage methodology to monitor and analyze candidate behavior during online examinations using video and audio inputs. Each stage is designed to process data efficiently and contribute to accurate detection of suspicious activities.

Initially, live video is captured from the candidate's webcam and divided into frames using a frame capture module. These frames are pre-processed using OpenCV techniques such as resizing, normalization, and noise reduction to ensure consistent input quality for further analysis.

The processed frames are then passed to the AI detection module, where multiple detection algorithms operate in parallel. The face detection and recognition module identifies the candidate using facial embeddings generated through deep learning models. Cosine similarity is used to compare the captured face with stored reference data for identity verification.

The gaze tracking module analyzes eye landmarks to determine the direction of the candidate's gaze. Eye Aspect Ratio (EAR) and pupil positioning techniques are used to detect attention shifts or prolonged deviation from the screen. Similarly, the head pose estimation module uses facial landmarks to calculate head orientation (yaw, pitch, and roll), helping identify abnormal movements.

For environmental monitoring, the object detection module utilizes a YOLO-based model to detect multiple persons or unauthorized objects within the frame. In parallel, the audio analysis module processes microphone input using MFCC (Mel-Frequency Cepstral Coefficients) features to detect speech or unusual background noise.

All detected outputs are combined in a decision layer, where rule-based thresholds and confidence scores are applied to determine violations. These events are logged with timestamps and used to generate a final report summarizing candidate behavior.

This multi-modal methodology improves detection accuracy by combining visual and audio cues while maintaining real-time performance.

VII. TECHNOLOGIES USED AND IMPLEMENTATIONS

The proposed online proctoring system is implemented using a combination of modern web technologies, machine learning frameworks, and computer vision libraries to ensure real-time performance and accuracy. The frontend of the system is developed using React.js, which provides a responsive and interactive user interface for candidates and examiners. It handles user authentication, exam interface rendering, and captures webcam and microphone input using WebRTC APIs.

The backend is implemented using Python, with frameworks such as FastAPI or Flask to manage server- side operations. The backend is responsible for processing video frames and audio streams, executing detection algorithms, and communicating results to the frontend in real time.

For computer vision tasks, OpenCV is used for frame processing and image manipulation, while MediaPipe and Dlib are utilized for facial landmark detection, gaze tracking, and head pose estimation. FaceNet or similar deep learning models are used for face recognition by generating embeddings and comparing them using cosine similarity. The system also incorporates a YOLO-based object detection model to identify multiple persons or unauthorized objects within the frame. For audio analysis, MFCC (Mel-Frequency Cepstral Coefficients) features are extracted and processed to detect speech or background noise.

A database system such as SQLite or PostgreSQL is used to store candidate details, violation logs, and session records. Additionally, WebSocket or REST APIs are used for real-time communication between the frontend and backend.
Overall, the integration of these technologies enables efficient real-time monitoring, accurate detection, and scalable deployment of the system.

VIII. RESULT AND DISCUSSION

The following figures represent the implementation and working interface of the system.

Fig.2. signup Page

Fig.3. Login Page

Fig.4. Exam

Fig.5. Proctoring Module

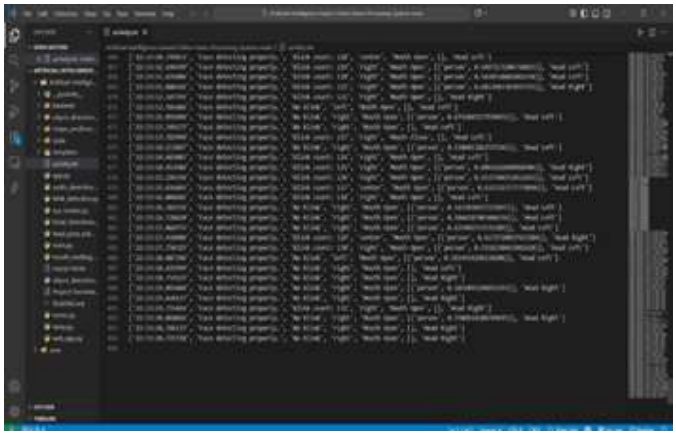

Fig.6. Activity Log

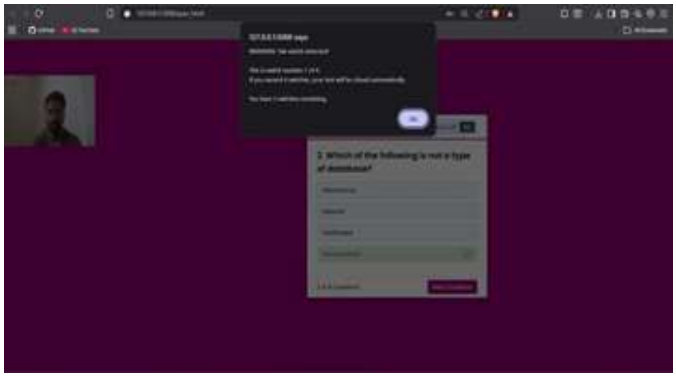

Fig.7. Tab Switch Detection

Fig.8. Test Result

The system was evaluated across multiple test sessions under varying lighting and network conditions. It achieved an overall accuracy of 92.6%, with precision of 87% and recall of 81%. The system maintained real-time performance at 20–30 FPS with an average latency of approximately 170 ms.

The results demonstrate that the system is capable of detecting suspicious activities effectively while maintaining low false positives.

IX. CONCLUSION

In this paper, we presented an AI-based online proctoring system designed to address the challenges of maintaining fairness and integrity in remote examinations. The system integrates multiple technologies, including face recognition, gaze tracking, head pose estimation, object detection, and audio analysis, to monitor candidate behavior in real time. By combining these different modalities, the system is able to detect a wide range of suspicious activities more effectively than traditional or single-method approaches.

During implementation and testing, we observed that the system achieved a high level of accuracy while maintaining real-time performance. The use of automated monitoring reduces the dependency on human invigilators and eliminates issues such as fatigue and inconsistency. Additionally, the modular design of the system allows it to be scalable and adaptable for large-scale online examinations.

Although the system performs well under standard conditions, certain limitations such as sensitivity to low lighting and camera quality were identified. These challenges highlight areas for further improvement.
Overall, the proposed system provides a reliable, efficient, and scalable solution for online exam monitoring. It contributes to enhancing the credibility of remote assessments and supports the growing need for secure digital examination environments.

X. FUTURE SCOPE

While the proposed system demonstrates effective performance in monitoring online examinations, there are several opportunities for further enhancement. One important direction is the integration of emotion and stress detection, which can help in understanding candidate behavior more deeply and identifying unusual patterns beyond basic movements.

The system can also be extended to support multi- camera monitoring, allowing detection of activities occurring outside the primary webcam view. This would help in preventing advanced cheating methods involving secondary devices or off-screen assistance.

Another potential improvement is the use of federated learning techniques, which would allow the model to improve over time without transferring sensitive user data to centralized servers, thereby enhancing privacy and security. Additionally, optimizing the system for low-light conditions and low-quality cameras will make it more accessible to users with varying hardware capabilities.
The integration of mobile device detection and browser activity tracking can further strengthen the system's ability to detect

malpractice. Moreover, seamless integration with existing Learning Management Systems (LMS) can improve usability and adoption in educational institutions.

Overall, future developments will focus on improving accuracy, robustness, privacy, and scalability, making the system more reliable for real-world large-scale deployment.

REFERENCES

1. Y. Atoum, L. Chen, A. X. Liu, S. D. H. Hsu, and X. Liu, "Automated Online Exam Proctoring," IEEE Transactions on Multimedia, vol. 19, no. 7, pp. 1609–1624, 2017.
2. A. J. Swart and K. Lombard, "Evaluation of Online Invigilation Techniques," in Proc. IEEE Global Engineering Education Conference (EDUCON), 2017, pp. 1794–1802.
3. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2015.
4. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE CVPR, 2016.
5. D. E. King, "Dlib-ML: A Machine Learning Toolkit," Journal of Machine Learning Research, vol. 10, pp. 1755–1758, 2009.
6. G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.
7. Google, "MediaPipe: A Framework for Building Perception Pipelines," 2020.
8. S. Davis and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 28, no. 4, pp. 357–366, 1980.
9. P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," in Proc. IEEE CVPR, 2001.
10. M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2016.
11. A. Nigam, R. Pasricha, and S. Singh, "A Systematic Review on AI-Based Online Proctoring Systems," IEEE Access, vol. 9, pp. 123–135, 2021.
12. H. M. Alessio, N. Malay, K. Maurer, S. Bailer, and J. Rubin, "Examining the Effect of Proctoring on Online Test Scores," Online Learning Journal, vol. 21, no. 1, pp. 146–161, 2017.
13. R. Bawarith et al., "E-exam Cheating Detection System," International Journal of Advanced Computer Science and Applications, vol. 8, no. 4, pp. 176–181, 2017.
14. S. Prathish, K. Bijlani, and T. S. Ashwin, "An Intelligent System for Online Exam Monitoring," in Proc. International Conference on Information Science (ICIS), 2016.
15. A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proc. NIPS, 2012.
16. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proc. ICLR, 2015.
17. R. Girshick, "Fast R-CNN," in Proc. IEEE ICCV, 2015.
18. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
19. European Union, "General Data Protection Regulation (GDPR)," 2018.
20. U.S.Department of Education, "Family Educational Rights and Priva