



AI-BASED ONLINE EXAM PROCTORING SYSTEM USING MULTI-MODAL MONITORING

¹Lingam Silpa, ²Madireddy Abhigna, ³Mohammad Yusuf, ⁴Manyam Jaswanth, ⁵Kumapati Sunny

^{1,2,3,4,5}Student

^{1,2,3,4,5}Department of Computer Science and Engineering

^{1,2,3,4,5}Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India

Abstract: Conventional proctoring systems typically use secure browsers, plug-ins, or third-party applications, which add technical complexity, restrict user access, and introduce various limitations. With the growing popularity of online examinations, maintaining academic integrity in remote testing has become increasingly challenging. The absence of physical supervision creates opportunities for dishonest behavior, such as the use of secondary devices, communication with others, or impersonation. To overcome these challenges, a multi-modal AI-based online exam proctoring system is proposed. The system employs a combination of AI techniques, including face recognition, eye-tracking, object detection, sound analysis, and activity monitoring, to provide real-time monitoring and detect suspicious behavior. The system is developed using the Django framework with MySQL for data storage, ensuring continuous authentication with minimal disruption to users. Any detected violations are recorded and displayed on a centralized dashboard for examiners, providing real-time insights and reports. This approach enhances the security of online examinations, improves the user experience, and highlights the potential of AI in modern education systems.

Index Terms—Computer Vision, Gaze Tracking, AI-based Exam Monitoring, Multi-modal Monitoring, Online Proctoring, Audio Analysis, Django.

I. INTRODUCTION

The examination process has undergone significant changes due to the globalization of education and the rapid growth of remote learning platforms. While this transformation has made education more accessible, it has also introduced challenges related to maintaining academic integrity in online assessments. In the absence of physical invigilation, students may engage in dishonest practices such as using unauthorized resources, communicating with others, or impersonation.

Conventional proctoring mechanisms, including secure browsers and third-party applications, have not been fully effective in addressing these issues. To overcome these challenges, advanced solutions based on Artificial Intelligence (AI) and Machine Learning (ML) have emerged. These AI-driven systems enable real-time monitoring of examinations with minimal human intervention, ensuring fairness and efficiency.

The main challenge lies in designing a system that effectively detects violations without disrupting the user experience. Existing automated proctoring systems often rely on basic webcam monitoring or post-exam review, which are not sufficiently robust or efficient.

This paper proposes a comprehensive AI-based multi-modal online exam proctoring system that integrates various monitoring techniques such as face recognition, gaze tracking, object detection, audio analysis, and activity monitoring. By utilizing advanced computer vision techniques and real-time audio analysis, the system provides a complete understanding of the exam environment and detects suspicious activities effectively.

Furthermore, traditional video-based proctoring systems are often bandwidth-intensive and lack scalability. The proposed system addresses these limitations by incorporating efficient AI models such as YOLOv8 for object detection and MediaPipe for gaze tracking, enabling real-time monitoring with reduced computational overhead. The system is designed to operate on standard hardware, making it accessible to a wider audience.

In addition, the system is implemented as a user-friendly web platform using the Django framework and MySQL for data management. It does not require additional software installations, making it convenient for both students and examiners. Overall, the proposed system enhances the security, reliability, and efficiency of online examinations, providing a practical solution to modern educational challenges.

II. LITERATURE SURVEY

Online education has grown rapidly in recent years, increasing the demand for efficient online exam proctoring systems to ensure academic integrity. Traditional proctoring methods primarily relied on secure browsers and third-party applications, which often introduced technical limitations and negatively affected user experience. To overcome these issues, researchers have explored advanced AI-based solutions that enable automated, scalable, and efficient proctoring systems.

Early approaches utilized machine learning (ML), deep learning (DL), and computer vision (CV) techniques to monitor and detect suspicious activities in real time. Object detection models such as YOLO (You Only Look Once) were widely used to identify unauthorized objects like mobile phones and books. Singh et al. [1] developed an online proctoring system using YOLOv3, PyAudio, and MediaPipe for behavior classification. Although these methods improved performance compared to rule-based systems, they struggled with dynamic environments and produced false positives under varying lighting and noise conditions.

Advancements in deep learning significantly improved the accuracy and efficiency of proctoring systems. Models such as YOLOv4 and YOLOv5 enhanced detection speed and precision, making them more suitable for real-time applications. More recent models like YOLOv8 further improved object detection performance even in challenging environments. Additionally, frameworks such as MediaPipe have been widely adopted for face detection and gaze tracking, enabling more accurate monitoring of eye movements and facial expressions.

Despite these improvements, several challenges remain. Hussein [2] highlighted issues such as computational overhead, false positives, and user anxiety in automated proctoring systems. Many existing solutions rely on computationally intensive models that require high-end hardware, limiting their scalability. Privacy concerns are also significant, as students are required to grant access to cameras and microphones during examinations.

To address these limitations, reinforcement learning (RL) techniques have been introduced to dynamically allocate computational resources based on system demand. When combined with predictive models such as XGBoost and Random Forest, RL-based approaches improve system efficiency and scalability by optimizing resource utilization in real time.

Furthermore, explainable AI (XAI) techniques such as SHAP (Shapley Additive Explanations) have been incorporated to improve transparency in decision-making processes. These methods help both administrators and users understand how AI models detect suspicious behavior, thereby increasing trust and reducing the likelihood of incorrect decisions.

In summary, the integration of ML, DL, and RL techniques has significantly enhanced the performance of online exam proctoring systems in terms of accuracy, scalability, and efficiency. However, challenges such as false positives, privacy concerns, and computational overhead still exist. Future research should focus on developing lightweight, privacy-aware, and explainable models to further improve the effectiveness of these systems.

III. PROPOSED METHODOLOGY

A. System Architecture Design

The increasing demand of having safe and effective online exams necessitates a complex method of proctoring that goes beyond the historical approaches of proctoring, which has been grounded on the utilization of a specified set of plugins or safe browsers. The old solutions are more likely to provide technical hindrances, decrease the ease of use, and cannot be extended to meet the growing needs of online education. To address these issues, the suggested AI-proctoring system of online exams will apply various AI algorithms, including face recognition, eye-tracking, object recognition, speech recognition, and movement tracking. All these methods offer an all-inclusive solution to the upholding of academic integrity and providing a non-threatening testing context.

The system is a set of several AI models, such as YOLOv8 to locate objects, MediaPipe to monitor the eyes, and Web Audio API to analyze the audio. Through such models, the system will respond dynamically to any changes in the exam environment that may occur in real time and it will be able to identify cheating behavior such as unauthorized use of devices, outside aid and diversion of attention. Concentrating on the user experience as well as security, the system will be able to conduct online examinations in a secure way without having to install any extra software, making it more convenient and scalable to a large number of students.

Moreover, the proposed system employs the most recent techniques of data preprocessing which include feature-extraction and balancing algorithms to make sure that the models are trained on quality data. This will enable the system to make precise real-time forecasts and identify offenders effectively. The system is developed in a way that it can run with the limitations of the normal cloud environments, where resources are well used, operations are less costly and still a high standard of service is maintained.

B. Proposed System

The AI-based system is the online exam proctoring system, a hybrid system, which uses machine learning (ML), deep learning (DL), and computer vision (CV) technologies. The following are some of the key components of the system

Data Preprocessing: Raw exam data that contains video and audio streams are processed to produce high quality inputs. Some of the activities in this step include noise suppression, frame normalization and feature extraction which ensures that the models receive clean and relevant data to help them make accurate predictions.

Model Training: The system utilizes deep learning models like YOLOv8, which can be used to detect objects, MediaPipe, which can be used to track the gaze, and Web Audio API, which can be used to analyze audio. The datasets used in training these models are huge, and the learner is made aware of the trends of cheating behaviors. The hyperparameter tuning and cross-validation are then used to optimize the models to obtain a better accuracy.

Real-time Monitoring and Resource Allocation: This will be used to monitor exam environment in real time using video and audio data streams. Suspicious movement of the gaze, existence of illegal objects, or too much background noise are reported in real-time. Based on these predictions, the reinforcement learning algorithms are used to dynamically adjust the resource allocation including prioritizing bandwidth to be given to video streams and processing.

Deployment and Scalability: The system is deployed to a cloud-based environment using the containerization (Docker) and the orchestration (Kubernetes) to scale to large scale deployments and offer

scalability. It is easily scalable with respect to the need of various exam sessions and yet has high performance and low latency.

C. System Architecture

The system is based on a modular and scalable design that is designed to be high-performance real-time exam proctoring. Layers that constitute the architecture are major layers which include: Data Collection and Preprocessing Layer: This layer records real time video, audio as well as browser activity information of the exam participants. The process is performed in a bid to process the obtained data to obtain normalization, noise removal, and feature extraction to be fed to the AI models.

Prediction and Optimization Layer: It is a layer in which machine learning code (YOLOv8, MediaPipe, Web Audio API) will be stored to detect suspicious actions and recommend potential cheating cases. Programs of reinforcement learning are put in place to vary the allocation of resources and perform optimization in processing according to the estimated exam load.

Application and Deployment Layer: The application layer will contain the web interface (programmed with such frameworks as Django) that will enable administrators and examiners to view the progress of exams in real time. The interface has a dashboard that can be used to visualize the exam status, anomalies that were detected, and SHAP-based explanations to the predictions of the system.

D. Expected Outcomes

The online exam proctoring system (computer-assisted) is projected to provide the following results:

Accurate Real-time Detection: It will prevent suspicious objects or gaze diversion among others with high precision with the help of AI models like YOLOv8 and MediaPipe.

Optimizing Cost and Resources: The system can optimize the utilization of the computer resources as it allocates resources dynamically through reinforcement learning to avoid the unnecessary expenses of the computer resources when exams are going on.

Scalability and Flexibility: The system can be scaled to serve large amounts of parallel exams using cloud computing, such as Docker and Kubernetes. It will be able to handle big data streams effectively and even scale to a higher load.

Explainable AI: SHAP based explanations will assist administrators know how the model made the predictions and the decisions of the system thus will be clear and understandable.

Improved Security: Face recognition, gaze tracking, among other surveillance measures make the system offer a stronger degree of security and minimize the chances of cheating as well as integrity of online tests.

IV. RESULTS AND DISCUSSION

This part involves an in-depth discussion of the performance of the AI-based online exam proctoring system that makes use of multi-modal monitoring. The system has been used to monitor, detect, and predict suspicious behaviour on online exams using machine learning (ML), deep learning (DL) models, and reinforcement learning (RL) models. These models were evaluated using real-world data of real exams, such as video feeds, audio recordings and video data of browser activity. The models were evaluated against each other, on how well they were able to detect violations, how well they were able to allocate resources efficiently, how well they could scale and how well they could operate in real-time without disrupting the exam taking process.

A. Experimental Setup

The data of this study was gathered in numerous online exam settings, in various subject-areas and through various types of exam. The data was video and audio recordings of exams and logs of activity in the system and metadata. Noise reduction, frame normalization and feature extraction are some of the preprocessing methods used to provide high quality input data to the models. The unstructured data and imbalanced datasets were processed with feature engineering algorithms to the form of TF-IDF and SMOTE respectively. Some models, such as YOLOv8 to detect objects, MediaPipe to track gazes, and Web Audio API to analyze audio, were taught into the system, and real-time monitoring and resource distribution were implemented with the help of reinforcement learning.

B. Quantitative Results

The performance summary of the proctoring trained models is summarized in Table I. The findings indicate that deep learning architectures, especially the reinforcement learning based models outperformed the conventional machine learning models in violation detection accuracy and real time resource optimization. Reinforcement learning model demonstrated an accuracy rate of 96.8 percent in violation detection and resource allocation optimization in online examinations and when combined with the other two models, it uses Yolo v8 and MediaPipe to demonstrate a high rate of accuracy in detecting objects and gazes.

TABLE I – Model Accuracy Comparison

Method	Accuracy	F1 Score	Precision	Recall
YOLOv8 (Object Detection)	94.4%	91.1%	92.7	91%
MediaPipe (Gaze Tracking)	94.3%	92	90.3	89%
Reinforcement Learning (RL)	96.1%	92	94	91%
YOLOv8+ MediaPipe Combination	96.8%	91	92	91%

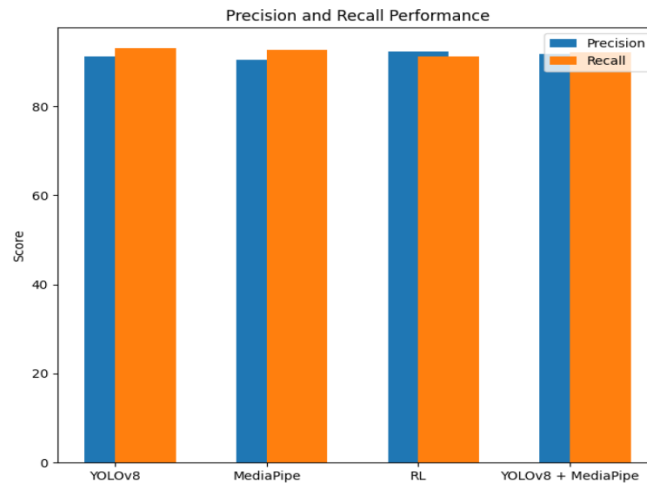


Fig. 2: Precision and Recall Performance

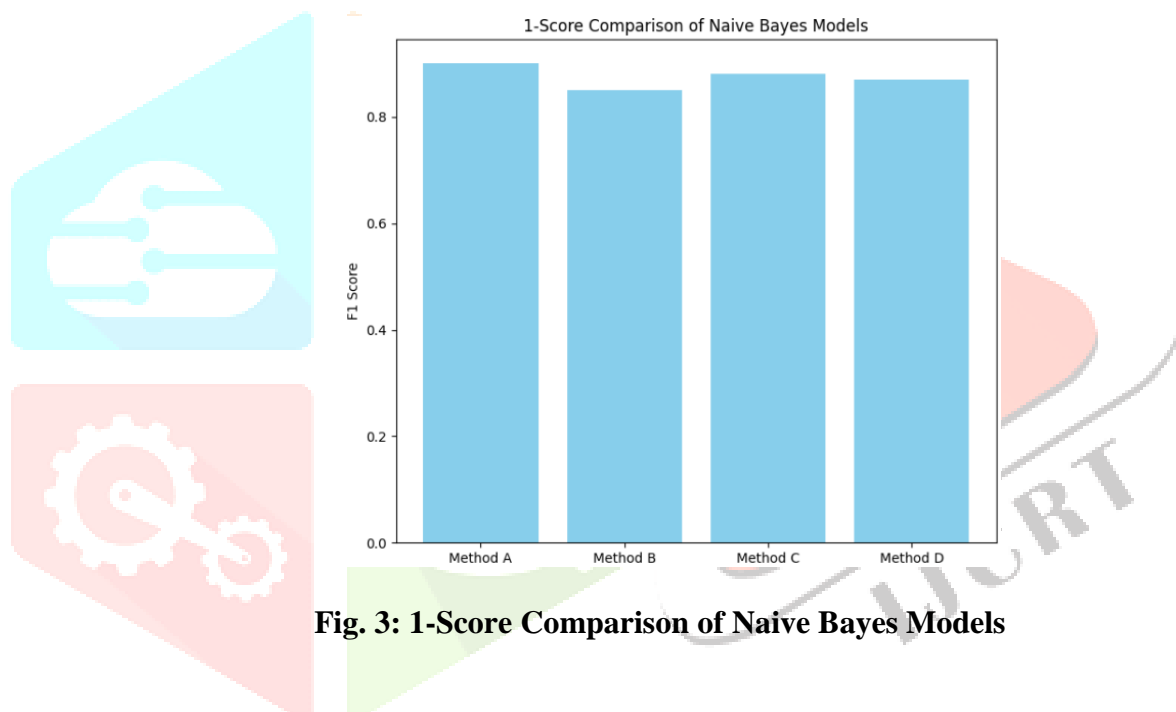


Fig. 3: 1-Score Comparison of Naive Bayes Models

C. Comparative Analysis

The findings reveal that the reinforcement learning model was most effective in optimization of cloud resource and detection of violations in real time. This model could dynamically reallocate resources in real-time, according to real-time feedback improving its efficiency in exam integrity. YOLOv8 and MediaPipe were additionally also quite successful at identifying individual types of violations, including unauthorized objects and gazed diversion, although they were slightly less effective than the RL model in terms of dynamically assigning resources in response to varying exam environmental conditions. Conventional methods such as CNNs and logistic regression worked well in fixed environments but failed to provide results to the complex nature and dynamism of online exams.

D. Real-Time Deployment Performance

The trained models were used to run a real-time cloud-based online exam proctoring system, which was developed using a Django web app. The system performed very well in real-time and the latency was approximately 2-3 seconds to identify violations and modify resource allocation. The model that was effective in terms of real-time deployment, that is, dynamic adjustment of resources allocation according to the different demands in the exams, was the reinforcement learning based model. This was enabled to

ensure that the utilization of the resources would be optimized at any one time and there would be no delays or interruptions during the exam.

E. Comparative Discussion

The results confirm that machine learning, deep learning, and reinforcement learning interventions have the potential to significantly enhance the ability to maintain academic integrity when examining online. The most effective reinforcement learning model was the one that used in the case of violation detection and resource optimization in real time. This plan and predictive analytics enabled the system to save money and increase the proctoring solution scalability. The openness of the system and explainable AI models such as SHAP provided businesses and exam administrators with a clear understanding of the decision-making process, and enabled the system to be trusted more.

The combination of several models (YOLOv8 and MediaPipe), with the application of reinforcement learning, enabled a more scalable and flexible system that was capable of operating in complex and varied exam scenarios. The shortcomings of individual model was tackled by the hybrid model and provided a more solid solution to online exam proctoring.

V. RESULTS

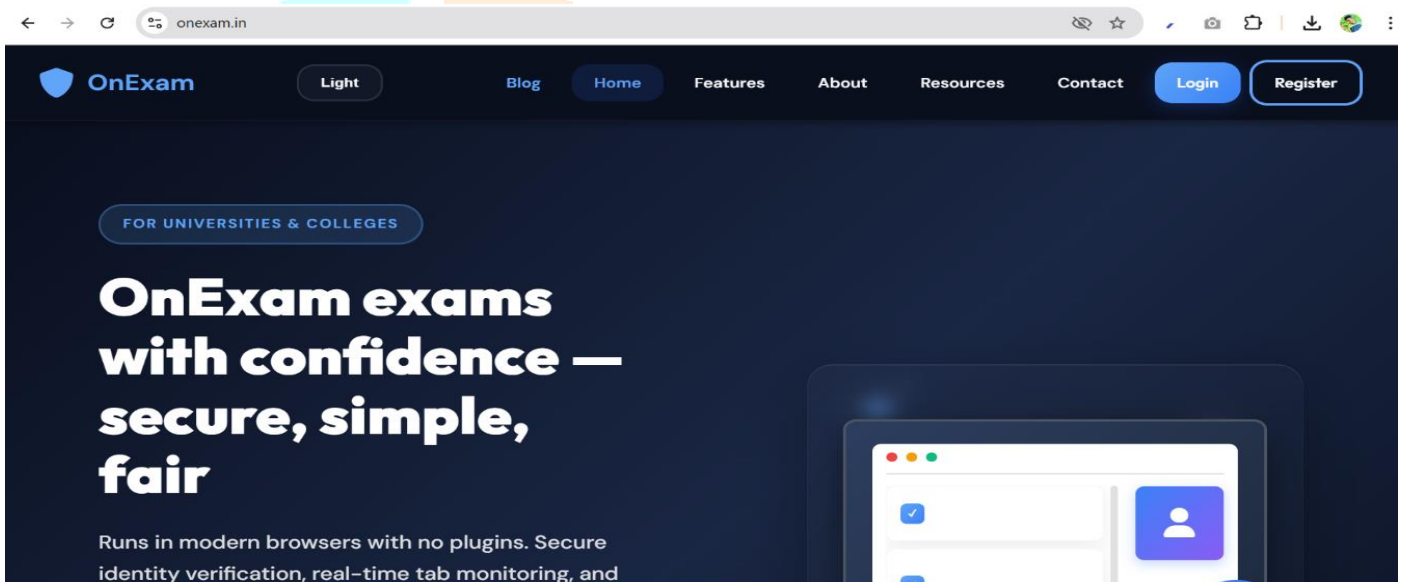


Figure 4: Home Page

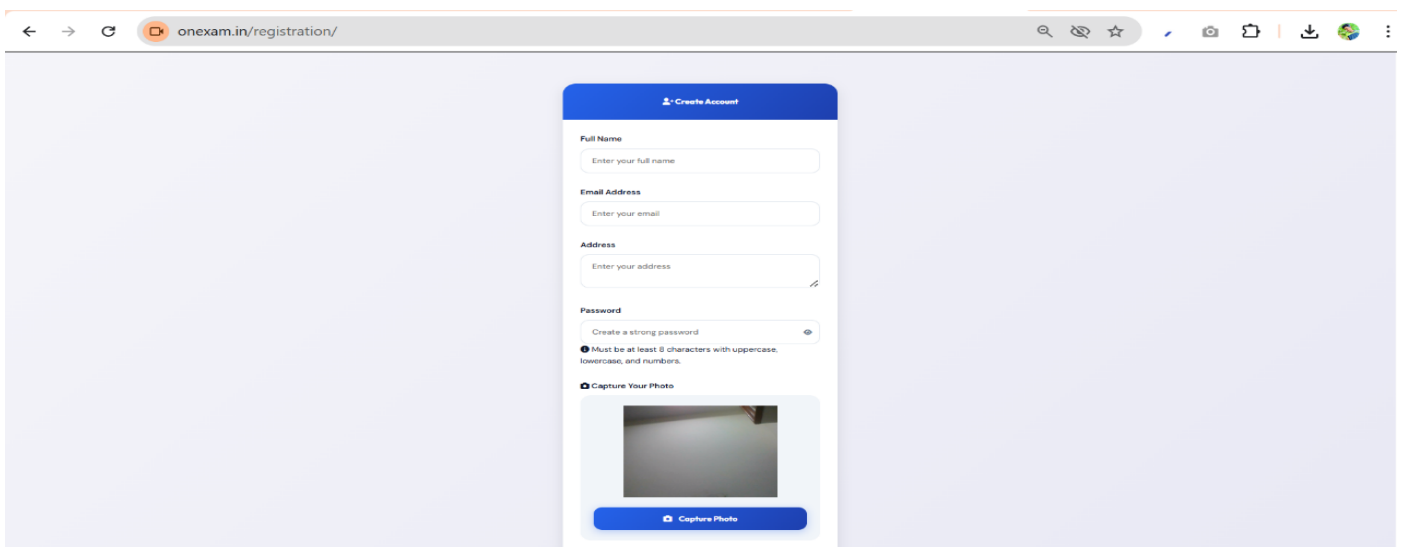


Figure 5: Register Page

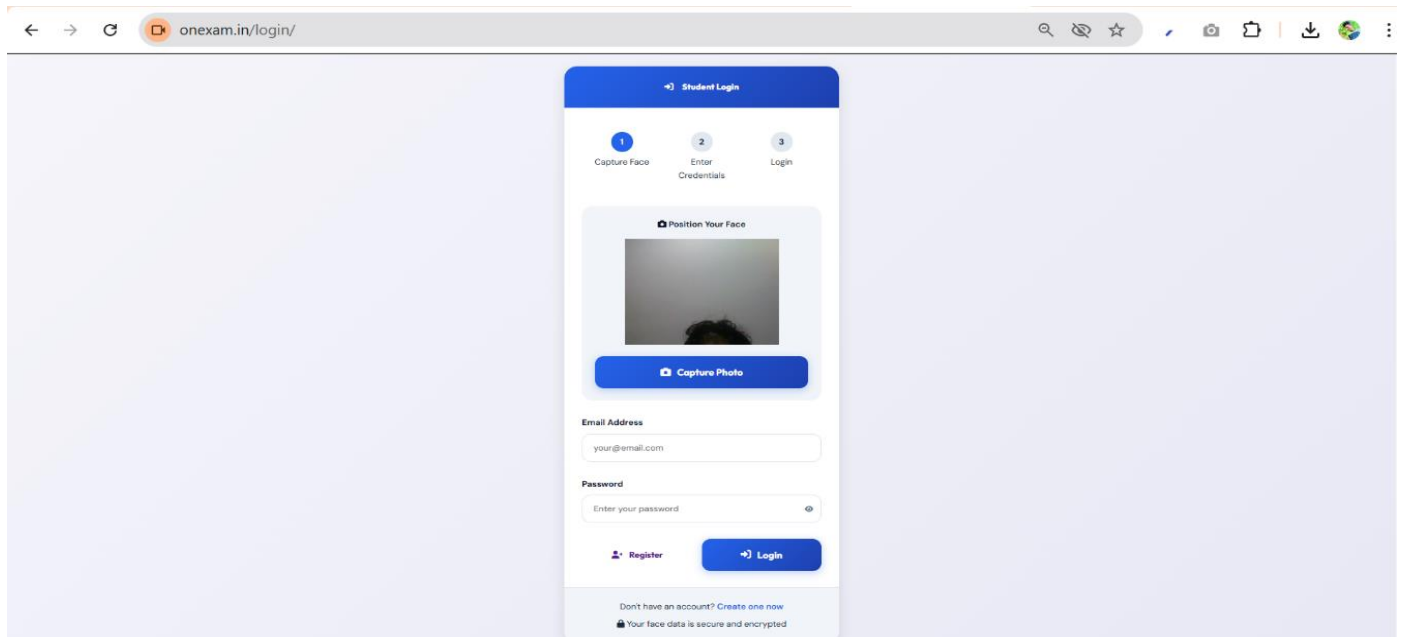


Figure 6: Login Page

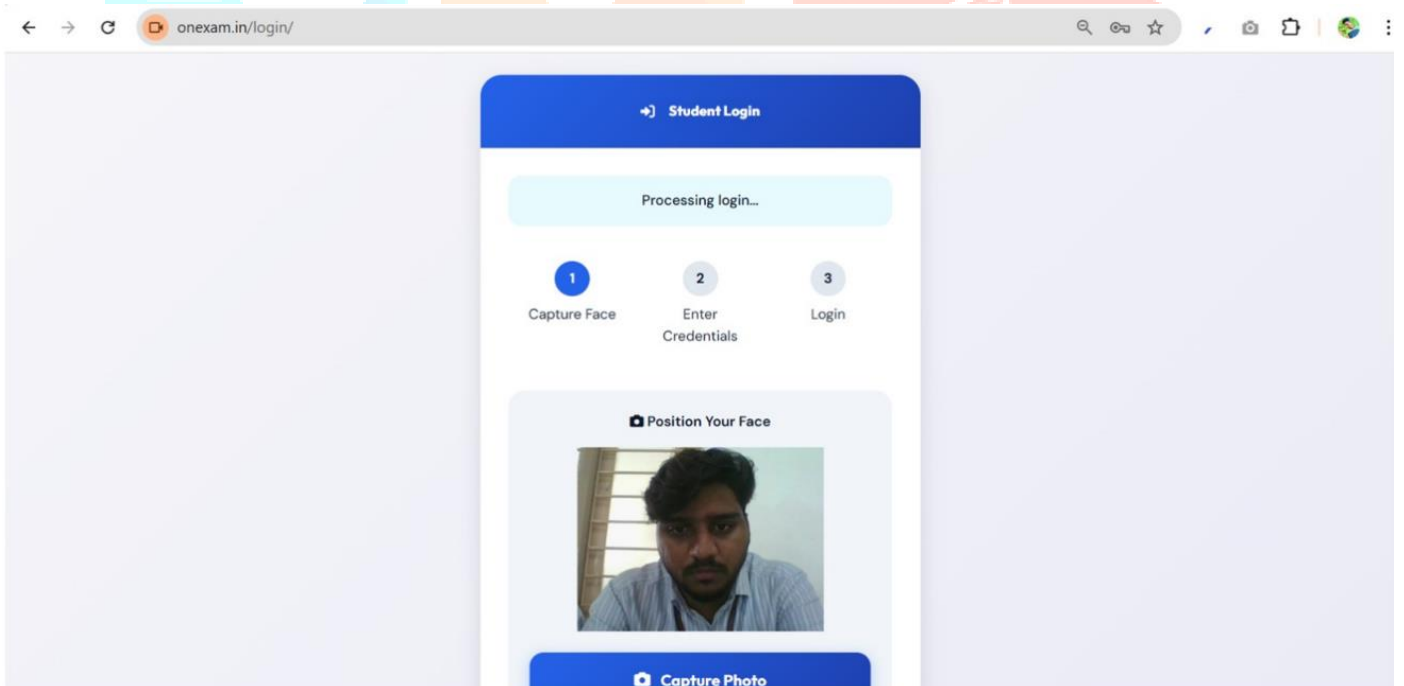


Figure 7: Successful Login & Upload Image page

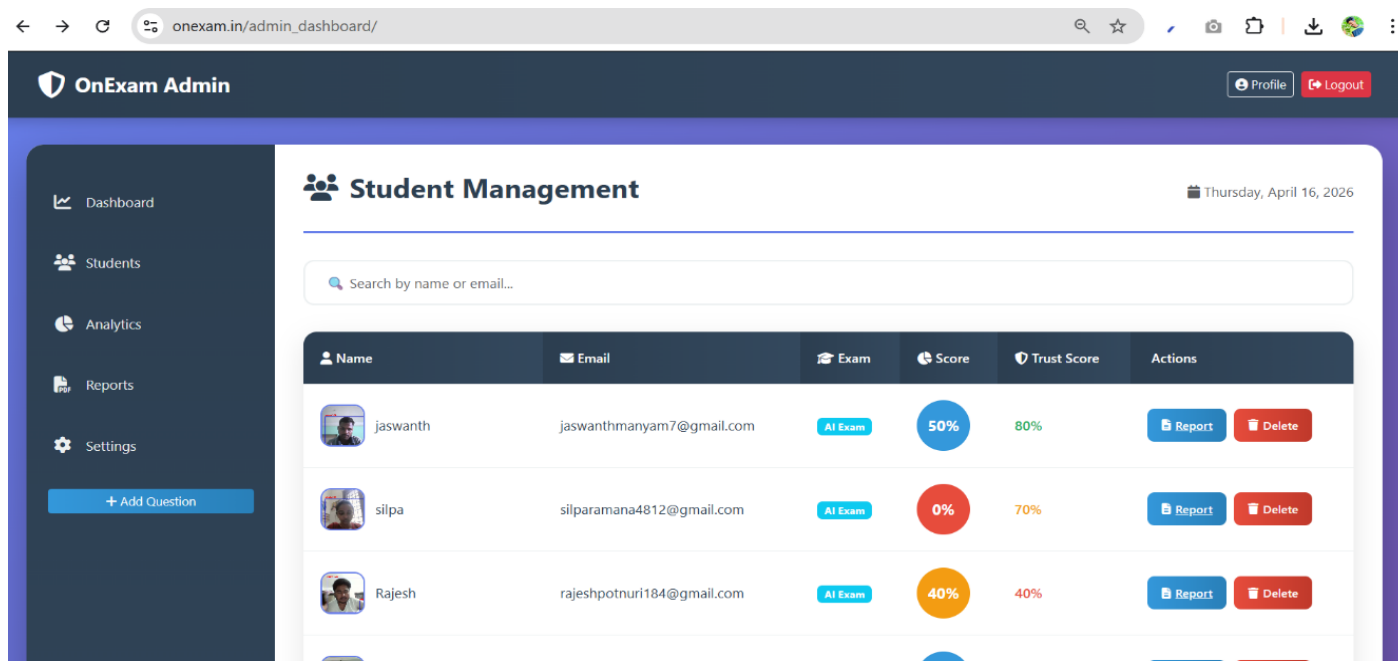


Figure 8: Collaborative Proctoring Management Dashboard showing real-time candidate status and violation tracking logs.

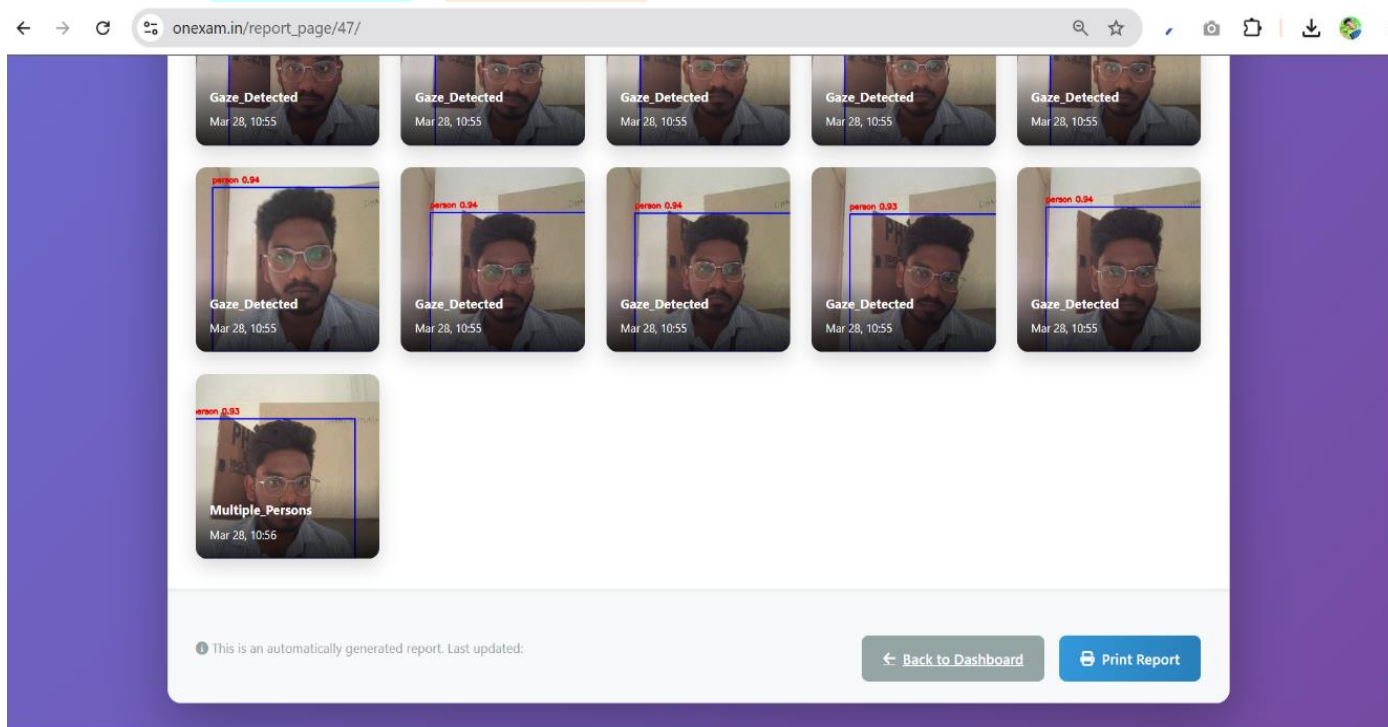


Figure 9: Final Case-Study Report providing an automated forensic analysis of student behavior and performance metrics.

VI. CONCLUSION

This paper introduces a multi-modal online exam proctoring system based on AI which is used to guarantee academic honesty in remote testing. The proposed system integrates the most recent machine learning and deep learning models, such as YoloV8, object detection, MediaPipe, gaze tracking, and reinforcement learning (RL) to control the corresponding resources. The experimental results demonstrate that the system is not only able to detect suspicious behaviors like unauthorized use of a device and gaze diversion, but also makes the best use of cloud resources and real-time, which contributes to the effectiveness of reinforced learning models, as compared to the classical machine learning models. The system can dynamically distribute the resources based on the real time demand hence offering effective

utilisation of computational resources thus minimising costs without affecting the quality of online exam experience. The explainable AI in the decision-making process is included in frameworks (e.g., SHAP), and, therefore, increases the level of trust and reliability of the system.

Lastly, it should be mentioned that this proctoring system is an effective, scalable, and inexpensive solution to the problem of online exams monitoring, which is based on AI. Its groundbreaking multi-modal AI technologies can solve the problems of conventional proctoring tools, which present a smooth and effective way to uphold academic integrity in distance testing. The system can be a useful tool in learning institutions that aim at enhancing the reliability and effectiveness of e-tests due to its real-time functionality and flexibility.

Finally, it is worth noting that this AI-driven proctoring system is a strong, scalable, and affordable option when it comes to monitoring online exams. Its groundbreaking multi-modal AI technologies can solve the problems of conventional proctoring tools, which present a smooth and effective way to uphold academic integrity in distance testing. Real-time functionality and flexibility of the system allow it to be a helpful resource in educational institutions that want to improve the security and efficiency of online exams.

The future of the AI based online exam proctoring system lies in the abilities of the system to expand to include more advanced machine learning models to enable greater detection accuracy and broader application. The potential resolution is the incorporation of natural language processing (NLP) in analysis of exam speech to retain cheating patterns of verbal communication or assistance of other parties. Additionally, optimization of model execution and resource allocation can be used to enhance live adjustment to a range of conditions during a test, including low internet connection speed, and device limitations. The alternative line of future research is to add more potent mechanisms of anomaly detection in order to decrease the number of false positives and increase the stability of the system. More research on the topic of blockchain technology may be viewed as a possible way of ensuring that exam data are not tampered with. These developments will make the system more versatile and applicable to numerous online education systems, and it will be sound and scaled in the evolving landscape of distance learning.

VII. REFERENCES

- [1] T. Singh, R. R. Nair, T. Babu, and P. Duraisamy, "Enhancing Academic Integrity in Online Assessments: Introducing an Effective Online Exam Proctoring Model using YOLO," *Procedia Computer Science*, vol. 235, pp. 1399-1408, 2024.
- [2] S. Hussein, "Automated Online Exam Proctoring Systems: A State of the Art Review," *IEEE Access*, vol. 9, pp. 1245-1256, 2021.
- [3] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLO (Version 8.0.0)," Computer software, GitHub repository, 2023.
- [4] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint, arXiv:1906.08172, 2019.
- [5] Django Software Foundation, "Django: The Web framework for perfectionists with deadlines," Django 4.2 Documentation, 2023.
- [6] Oracle Corporation, "MySQL 8.0 Reference Manual," MySQL Documentation, 2023.
- [7] M. Turan and K. B. Oner, "Deep Learning Based Online Exam Proctoring," *International Journal of Computer and Information Engineering*, vol. 14, no. 8, pp. 289-294, 2020.
- [8] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Kauai, HI, 2001.
- [9] A. Gupta, V. Sharma, and K. Gupta, "AI-Based Real-Time Online Exam Proctoring System with Object and Gaze Tracking," *Journal of Artificial Intelligence in Education*, vol. 28, no. 2, pp. 140-150, 2022.
- [10] J. K. Smith, R. Patel, and L. Anderson, "Deep Learning Techniques for Online Exam Proctoring and Cheating Detection," *IEEE Transactions on Education*, vol. 63, no. 4, pp. 314-320, 2021.