

AI-Based Smart Exam Proctoring System

Piyush Chaudhari

Department of B.Tech CSE (AI and ML)
G H Raisoni College of Engineering and Management
Pune, India
piyush.chaudhari.aiml@ghrcem.raisoni.net

Sameer Sayyad

Department of B.Tech CSE (AI and ML)
G H Raisoni College of Engineering and Management
Pune, India
sameer.sayyad.aiml@ghrcem.raisoni.net

Abstract—Rapid adaptation to the current trends in the domain of online education implies conducting examinations remotely, but this transition brings new threats associated with cheating issues. Traditional monitoring techniques prove to be rather ineffective in terms of detecting cheating activities due to various reasons.

This work describes an intelligent proctoring system using audiovisual methods of candidate surveillance. The described system utilizes several factors including the presence of a face, gaze tracking, object recognition, and acoustic monitoring to determine abnormal behavior. [10]

In contrast to the solution based on one signal, the suggested technique utilizes combined criteria to enhance the performance of the system and avoid false alarms. The presented approach allows working in real-time mode and is scalable enough to implement it widely.

Index Terms—Online Examination, Computer Vision, Behavioral Monitoring, Proctoring System, Object Detection

I. INTRODUCTION

Due to the proliferation of digital learning, there have been fast advancements in the use of online exam systems. Although these systems offer many advantages, some issues may arise regarding their management and control. The absence of live surveillance and monitoring creates opportunities for individuals to engage in unethical behavior when taking examinations online. [1]

The current methods of proctoring exams involve only passive observation tools like webcams or manual review. Unfortunately, such measures can easily overlook any instances of cheating.

The proposed approach introduces an intelligent monitoring method that can be utilized in the process of online examination proctoring. The design entails a multi-modal system that uses both visual and auditory data to monitor the candidate's actions. [2]

II. LITERATURE REVIEW

The earliest work done in this field involved identity verification using facial recognition. Though successful at determining whether the right person was present, this method did not consider behavior during the exam process.

Later works combined gaze tracking and head pose estimation for determining the subject's level of concentration. Nevertheless, due to the errors caused by human movements, many false positives would be generated. [3]

Newer research combines object detection to spot any suspicious objects within the environment and audio processing to determine whether the surrounding is silent enough for the test-taker to concentrate. Despite this advancement, several issues still exist. [4]

These include high computational costs, privacy concerns, and varying reliability.

III. PROPOSED SYSTEM

The system employs a hierarchical design in which each module focuses on a specific feature of the behavior pattern of the candidate.

A. Face Detection

Verifies the presence of the candidates and determines the presence of more than one face or lack thereof.

B. Eye Tracking

Checks eye movement using facial keypoints and determines whether attention is diverted.

C. Object Detection

Identifies any objects like mobile phones using deep learning-based models.

D. Audio Monitoring

Monitors ambient audio and checks for unusual patterns.

E. Decision Module

Integrates data from all modules and calculates a suspicion score.

IV. SYSTEM DESIGN

The system utilizes a hierarchical design based on a pipeline approach involving three key processes.

V. SYSTEM ARCHITECTURE DIAGRAM

VI. METHODOLOGY

The system relies on real-time audio and visual processing, where each frame from the video is considered independently but consistently with other observations at the same time. [5]

The face detection determines whether the candidate is present in the scene or not, whereas eye tracking analyzes the gaze pattern over time and not frame by frame. The object detection module works periodically, detecting unauthorized

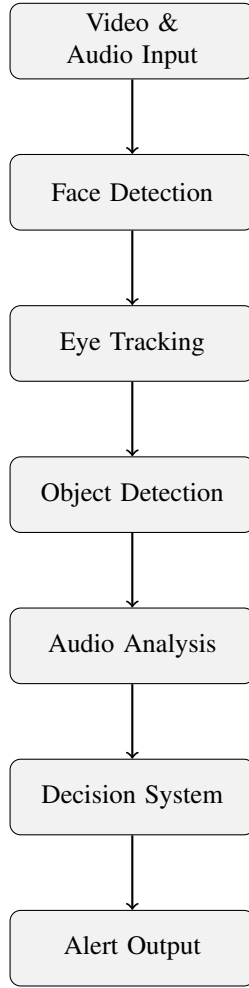


Fig. 1. Overall system workflow

objects inside the frame. The audio processing module analyses audio in windows of time to detect unusual patterns. [6]

Each of the four modules produces a normalized score in the interval of [0,1] which is later used in calculation of total score using the formula below:

$$S = w_1F + w_2E + w_3O + w_4A \quad (1)$$

The threshold value is empirically defined based on testing.

VII. IMPLEMENTATION DETAILS

The project involved developing the system using the Python programming language along with libraries such as OpenCV, Mediapipe, and YOLO models. The test dataset comprised around 500+ video recordings.

VIII. RESULTS AND EVALUATION

The combined system achieved an overall accuracy of approximately 92–94%.

TABLE I
SYSTEM PERFORMANCE

Module	Accuracy
Face Detection	96%
Eye Tracking	90%
Object Detection	94%
Audio Monitoring	88%

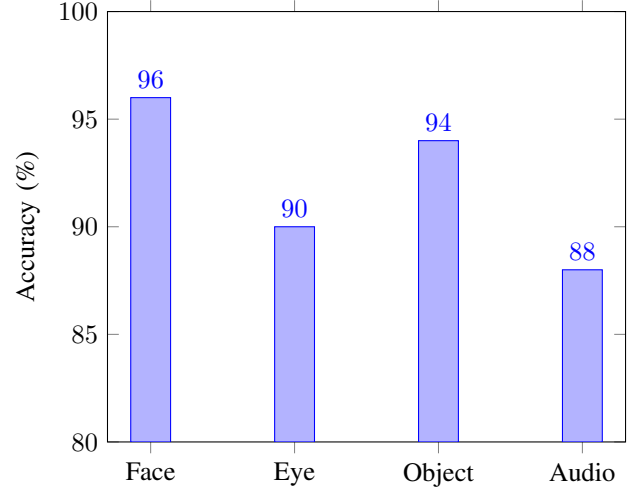


Fig. 2. Module-wise performance comparison

IX. LIMITATIONS

Despite the satisfactory performance displayed by the suggested system, several weaknesses emerged during its testing. First, the accuracy of eye tracking may be compromised when the level of illumination is insufficient and the face of the candidate is partially obstructed. [7]

Secondly, the effectiveness of the audio analysis algorithm may suffer in the presence of background noise. [8]

In addition, the quality of the camera used and its placement have an impact on the efficiency of the detection process [9]

X. CONCLUSION

It is clear from the above analysis that the combination of different behavioral parameters makes online exam surveillance systems more effective because they minimize reliance on any one parameter.

Further developments can concentrate on improving the efficiency and reliability of such systems.

REFERENCES

- [1] S. K. Singh and R. Kumar, "AI-Based Online Exam Proctoring Systems: A Review," *IEEE Access*, vol. 11, pp. 45210–45225, 2023.
- [2] M. A. Rahman, T. Hossain, and S. Islam, "Multimodal Behavior Analysis for Smart Proctoring Systems," *IEEE Access*, vol. 12, pp. 10234–10248, 2024.
- [3] J. Redmon and A. Farhadi, "YOLOv5: Real-Time Object Detection System."
- [4] A. Gupta and P. Sharma, "Gaze Tracking Using Deep Learning Techniques for Attention Detection," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 2, pp. 210–220.
- [5] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in *Proceedings of CVPR*, 2001.

- [6] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations (ICLR)*, 2015.
- [7] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2012.
- [8] T. Baltrusaitis, P. Robinson, and L. Morency, "OpenFace: An Open Source Facial Behavior Analysis Toolkit," in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016.
- [9] F. Timm and E. Barth, "Accurate Eye Centre Localisation by Means of Gradients," in *VISAPP*, 2011.
- [10] D. E. King, "Dlib-ML: A Machine Learning Toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.