International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Smart Exam Surveillance

Aditya Chauhan, Bhavya Jalan, Shrey Parashar, Dr. Anand Gupta

Department of Computer Science Netaji Subhas University of Technology Dwarka, Delhi 110028, New Delhi, India {adydelhi2003, shreyparashar44, Bhavyajalan.work}@gmail.com

Department of Computer Science Netaji Subhas University of Technology Dwarka, New Delhi 110078, India anand.gupta@nsit.ac.in

ABSTRACT—

The rapid rise of online learning has brought for- ward significant challenges in ensuring fairness during remote examinations. Traditional invigilation methods often fall short when it comes to identifying subtle or non-obvious forms of cheating, such as frequent off-screen glances or interactions with unauthorized tools. This project introduces an intelligent, automated proctoring system that leverages machine learn- ing—specifically pose estimation and eye-tracking—to contin- uously observe and assess candidate behavior in real time. Utilizing Convolutional Neural Networks (CNNs) in combination with tools like MediaPipe, the system tracks facial orientation and body posture to recognize unusual or suspicious activity, including persistent head movement or extended eye diversion. Such behaviors are flagged for further analysis, allowing timely alerts or interventions. Designed with scalability and minimal manual oversight in mind, the system provides a reliable and efficient way to uphold academic honesty in virtual exam settings. Furthermore, it establishes a foundation for future improvements and adaptability across diverse online assessment platforms.

***Index Terms*—Cybersecurity, XAI, Intrusion Detection System, SHAP, LIME, Ensemble Learning, Stacking, KDD Cup 1999.**

1. Introduction

A. Overview and Background

Examinations are a cornerstone of the education system, acting as key tools to measure students' understanding, skills, and knowledge in specific subjects. They play a critical role in assessing academic progress and ensuring that students meet established educational benchmarks. However, the credibility of examinations relies heavily on the principles of fairness and integrity. A fair assessment ensures equal opportunities for all students, reinforcing trust in the system and preserving the value of academic achievements. Instances of malpractice, such as cheating or unethical conduct, undermine the purpose of examinations and devalue the efforts of honest learners. Therefore, upholding fairness and integrity in examinations is crucial for both individual success and the broader educational ecosystem, which is built on the pillars of merit and ethical learning practices.

With the rapid expansion of online education, the need for secure and reliable remote proctoring solutions has become increasingly urgent. The growing adoption of online courses, degree programs, and virtual classrooms has led to a significant increase in online examinations. However, ensuring the in- tegrity of these exams in the absence of in-person invigilation presents unique challenges. Traditional proctoring methods, such as live video monitoring or facial recognition software,often fail to deter cheating effectively. These approaches can be prone to inaccuracies, generating false positives or negatives, and may also be perceived as overly invasive due to the need for constant observation. This project aims to address these limitations by leveraging machine learning (ML) techniques to detect pose and eye movements, enabling the identification of suspicious behavior during exams. Using a standard webcam, the system offers real-time monitoring of students' actions while safeguarding their privacy and eliminating the need for additional hardware. By focusing on key body parts such as the head, shoulders, and arms, the system provides a more comprehensive understanding of students' behavior during online tests. The solution employs multiple pose estimation models and eye-tracking techniques to monitor students in real time. It is designed to detect behaviors such as looking away from the screen, leaving the camera frame, or identifying the presence of unauthorized individuals, which may indicate cheating. Alerts are generated when such behaviors are detected, allowing invigilators to in- tervene as necessary. In essence, this project seeks to enhance the fairness and integrity of online examinations by integrating pose detection with eye movement analysis to identify and flag suspicious activities effectively.

B. Motivation and Rationale

The transition to online education has revolutionized the way exams are conducted, especially in the wake of global events such as the COVID-19 pandemic. Online learning platforms have provided unprecedented opportunities for ac- cessibility and scalability, yet they have also introduced critical challenges, particularly in ensuring the fairness and integrity of remote assessments. Traditional proctoring methods, such as manual supervision via video conferencing or automated systems relying on basic facial recognition and activity mon- itoring, have proven to be either intrusive, resourceintensive, or prone to inaccuracies. These limitations often result in unnecessary stress for students and increase the likelihood of false positives, eroding trust in the assessment process. This project is driven by the need to address these gaps and develop a more effective and student-friendly solution for online proc- toring. By harnessing advanced machine learning techniques, such as pose estimation and eye-tracking technologies, the aim is to create a robust system capable of monitoring student behavior in real time. Unlike conventional methods, this approach seeks to minimize false alarms, reduce invasiveness, and accurately interpret student actions during exams, striking a balance between fairness and privacy. Moreover, the global reliance on remote and hybrid learning models necessitates scalable and efficient solutions. Exams are not merely a test of knowledge but also a demonstration of fairness and equality. A reliable system that identifies suspicious behaviors—such as looking off screen, interacting with unauthorized devices, or leaving the exam area—while respecting students' privacy is essential to address the shortcomings of current proctoring practices. This project represents an innovative step forward in safeguarding academic integrity during online assessments. By incorporating advanced technologies, the system aims to empower educators and institutions with a reliable tool to uphold fairness while offering students a less stressful exam experience. The long-term vision is to establish a new stan- dard for proctoring that aligns seamlessly with the evolving landscape of online education.

C. Persistent Challenges in Online Proctoring

While significant progress has been made in automated remote proctoring, several key challenges continue to hinder the development of an accurate, scalable, and nonintrusive system: Accuracy and Minimization of False Positives De- tecting cheating behavior with precision remains a complex issue. Traditional systems often misclassify routine student actions—such as posture adjustments or brief distractions—as suspicious conduct. Improving classification accuracy while reducing false positives is critical to maintaining system cred- ibility. Limited Camera Field of View The reliance on standard webcam input restricts the observable area, leading to blind spots where students can potentially engage in unauthorized activities, especially when hands or materials are outside the camera's view. Hardware Constraints and Environmen- tal Variability Differences in hardware specifications (e.g., low-resolution webcams) and environmental conditions (e.g., poor lighting or cluttered backgrounds) can negatively impact the performance of visual tracking models used in remote proctoring. Real-Time Processing Requirements Live analysis of video streams for posture and behavior cues demands high computational power. Ensuring smooth operation with minimal latency while avoiding undue strain on users' de- vices is a significant technical challenge. Privacy and Ethical Considerations Persistent video monitoring, combined with facial and pose analysis, raises ethical and privacy concerns. Students may perceive such surveillance as overly intrusive, potentially affecting their performance or willingness to par- ticipate. Adaptability Across Diverse Environments Online exams are conducted in varied settings, from private homes to shared spaces. Designing a system that generalizes effectively across different environments, devices, and behavioral patterns without frequent retraining is crucial for widespread adoption.

D. Problem Definition

2. Literature Review

With the rapid adoption of online education, ensuring the fairness and authenticity of remote examinations has be- come a pressing challenge. Various technological interven- tions—ranging from facial recognition and gaze tracking to advanced pose estimation—have been developed to address this issue. This section critically explores the five evolutions of proctoring technologies, highlights innovations in behavior modeling, and examines how modern machine learning frame- works are reshaping the landscape of cheating detection.

A. Limitations of Early Proctoring Systems

Initial remote proctoring solutions relied heavily on webcam-based facial analysis and eye-tracking to monitor test- takers. These systems detected head movements and gaze deviations as indicators of suspicious behavior. However, studies such as Wang et al. [1] highlighted their limitations in detecting more complex behavioral patterns.

B. Temporal Behavior Modeling Through Action Recognition

To overcome the static nature of earlier approaches, re- searchers began employing action recognition models that analyze motion over time. These models use temporal data to distinguish between benign movements and deliberate attempts to cheat. For example, Li et al. [2] proposed spatio-temporal graph convolutions to recognize skeletal actions more accu- rately.

C. Pose Estimation as a Richer Behavioral Signal

Pose detection offers a broader and more informative behav- ioral signal than traditional facial tracking. Unlike systems that monitor only the face, pose estimation frameworks analyze full-body landmarks—such as the head, shoulders, arms, and torso—to detect nuanced physical cues. OpenPose by Cao et al. [6] is a significant contribution in this space.

D. Advancements in Deep Learning for Behavior Classifica- tion

Deep learning models have significantly improved the ac- curacy of automated behavior classification in proctoring sys- tems. Techniques like AutoEncoders and Transformers can detect complex behavior patterns from extensive pose and gaze datasets. These models use attention mechanisms to prioritize critical areas—such as the hands and face—for focused analysis [4].

E. Fusion of Gaze Tracking and Body Pose Analysis

Combining gaze tracking with body pose estimation enables a comprehensive view of user behavior. Since eye movement plays a vital role in understanding intention and engagement, this fusion allows the system to detect subtle, yet telling, signs of potential misconduct. Research by Dilini et al. [5] supports this integrated approach.

F. Relevant Datasets

Open-source datasets such as EX-Guard [8] and Cheating- VF [9] are increasingly used to train and evaluate cheating detection models in online proctoring environments.

3. Methodology and Experimental Setup

A. System Overview

This project presents an automated, non-intrusive online proctoring system that uses machine learning to monitor can- didate behavior. The system integrates Convolutional Neural Networks (CNNs) for facial analysis and MediaPipe Pose for body landmark detection, using only a standard webcam to ensure scalability, accessibility, and minimal hardware require- ments.

B. Convolutional Neural Networks (CNNs)

CNNs are used to detect facial behaviors such as eye clo- sure, mouth movement, and head orientation. Key advantages include:

- **Automatic Feature Extraction:** Learns relevant features without manual input.
- **High Accuracy:** Transfer learning enhances perfor- mance.
- **Real-Time Efficiency:** Optimized for fast inference.

CNN architecture includes convolutional, ReLU activation, and max-pooling layers. The output is passed through fully connected layers, ending in a sigmoid output neuron for binary classification.

C. Pose Estimation with MediaPipe

MediaPipe Pose detects 33 body landmarks from an RGB webcam feed. Benefits include:

- **Real-Time Processing:** Low-latency on-device inference.
- **High Precision:** Accurate across lighting and posture variations.

D. Key Algorithms and Concepts

1) Activation Functions:

- **Sigmoid:** Maps output to probabilities (0–1).

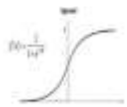

Figure 2: Visualization of Sigmoid Activation Function

- **ReLU:** Enables fast computation and deeper learning.

2) Backpropagation: Gradient descent is used to minimize Mean Squared Error (MSE) by updating weights and biases iteratively.

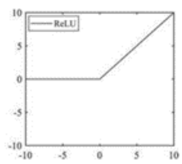

Figure 3: Visualization of ReLU Activation Function

3) Pooling: Max pooling reduces spatial dimensions, pre- serves important features, and enhances translation invariance.

E. Data Acquisition

Training uses two public datasets and user-generated videos with around 4,933 labeled images. Data is divided as:

- **Training:** Model learning phase.
- **Validation:** Avoid overfitting.

- **Testing:** Evaluate generalization.

F. Model Architecture

- **Input:** 150x150 RGB images.
- **Conv Block 1:** 32 filters, ReLU, max pooling.
- **Conv Block 2:** 64 filters, ReLU, max pooling.
- **Conv Block 3:** 128 filters, ReLU, max pooling.
- **Flatten:** Converts output to 1D.
- **Dense Layer:** 512 neurons, ReLU, dropout (0.5).
- **Output:** Sigmoid activation for binary classification.

G. Implementation Workflow

A. Data Acquisition

- Webcam feed is captured.
- Frames resized and normalized.

B. Feature Detection (MediaPipe)

- 468 face landmarks detected.
- Body posture tracked for suspicious behavior.

C. CNN Classification

- Frames passed through CNN.
- Output classifies normal vs. suspicious activity.

H. Tools and Frameworks

- **MediaPipe:** Real-time pose detection.
- **Keras:** Model construction and training.
- **OpenCV:** Webcam integration.

I. CNN Training

- **Dataset:** Labeled faces and behaviors.
- **Augmentation:** Zoom, rotation, flipping.
- **Optimization:** ReLU + Adam optimizer.

4. Results and Discussion

A. Conclusion

In the digital age, online exam proctoring technologies have the power to revolutionize the way assessments are admin- istered and tracked. These systems can identify applicants and track their behavior during the test using machine learn- ing, which enables real-time detection and disincentives for cheating. This guarantees that test results accurately represent a student’s knowledge and skills in addition to upholding academic integrity. This system’s non-intrusiveness is one of its main advantages since it provides a reliable and safe means of exam monitoring without violating test-takers’ privacy. All 15 things considered, our research shows how contemporary technology may facilitate trustworthy and equitable online tests, which is crucial given how quickly the educational scene is changing today. Online testing is inevitably becoming the next big concern with the growth of online learning, particularly in light of the difficulties caused by the COVID-19 pandemic. Although no system can be completely infallible, our project is an earnest attempt to provide a clever and workable solution to that increasing demand.

Figure 4: Overview of the CNN Architecture Used in the System

Fig. 1: ReLU Activation Visualization

B. Task, Achievement, and Possible Beneficiaries

This project's main goal was to build and create an au- tomated proctoring system that could use stance detection and eye movement tracking to keep an eye on students' behavior during online tests. Important behaviors that were effectively monitored and identified as possible signs of malpractice included frequent head movements, gaze devi- ations, and off-frame posture. This system's success rests in its capacity to operate with minimal human oversight while still detecting suspicious activity with a high degree of accuracy. It provides a scalable and affordable substitute for conventional invigilation techniques. Several players in the education ecosystem are among the benefits of this work. Universities and other educational organizations can use this approach to uphold exam standards in virtual environments with little opportunity for in-person supervision. It can be incorporated by e-learning providers and online testing plat- forms as a safe and trustworthy assessment tool, boosting user confidence. Because the approach encourages impartial and equitable review without being unduly intrusive, students also gain from it. The system promotes fairness and academic integrity by allowing real-time monitoring while maintaining user convenience. Such solutions have the potential to improve online education's integrity and reputation over time We implemented the framework in Python using common ML libraries (Scikit-learn for preprocessing and classical models, TensorFlow/Keras for the DNN, XGBoost for GBM, and the shap and lime packages for explanations). Key implementation details include: 1) Data Pipeline: The raw KDD-99 CSV files were parsed with Pandas. Categorical features were encoded using OneHotEncoder, resulting in a feature vector of dimension 120 (after encoding). Features were normalized to zero mean/unit variance. We discarded redundant features identified by a 0.99 correlation threshold. This preprocessing reduced training time and improved model clarity. 2) Model Training: The DNN was trained for 20 epochs with a batch size of 512, using Adam optimizer. Training converged quickly (training accuracy 98large dataset. The RF and GBM models were trained with default hyperparameters, but we also tuned the GBM's learning rate (finding 0.1 optimal). The stacking metamodel was fitted on the validation subset of predictions. Overall, the ensemble achieved 92set, with F1-scores above 0.90 on both classes (attack vs normal). These results match or exceed prior work on KDD-99. For example, Mane and Rao (2021) reported about 82improved performance likely stems from ensembling and updated preprocessing.

Future Work

The project holds significant potential for further advance- ments to enhance its efficiency and broaden its applications:

1) **Audio and Environmental Context Monitoring:** The system's ability to identify questionable sounds during tests, including conversations or gadget usage, can be improved using audio analysis. By integrating pose de- tection with ambient context monitoring, the proctoring system can offer a more comprehensive approach to preserving exam integrity.

2) **Training on Larger and Diverse Datasets:** The re- silience of the model can be enhanced by enlarging the training dataset to include various postures, cul- tural differences, and cheating scenarios. This expansion would allow the system to adapt to a wider variety of behaviors and environmental factors, thereby increasing its accuracy across diverse user groups.

3) **Enhancement of Pose Detection Accuracy:** The ac- curacy of the pose detection system can be improved through the adoption of more advanced deep learning models, such as transformer-based architectures or hy- brid models. This would address challenges such as var- ied camera angles, low lighting conditions, and partial occlusions, thus enhancing reliability in real-world exam scenarios.

4) **Privacy-Preserving Approaches:** Future versions could incorporate more privacy-conscious designs, such as anonymizing user data or effectively masking faces during recordings, to address ethical concerns related to monitoring.

5) **Generalization Beyond Exam Proctoring:** The pro- posed approach can be extended to other domains, including healthcare (e.g., identifying abnormal patient postures), workplace productivity monitoring, and driver monitoring systems for detecting fatigue or inattentive- ness.

References

[1] J. Wang, Z. Zhang, and H. Yu, "Combining CNNs with LSTMs for Human Action Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.

[2] C. Li, Z. Cui, W. Zheng, C. Xu, and J. Yang, "Spatio-Temporal Graph Convolution for Skeleton Based Action Recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, 2018. doi: 10. 1609/aaai.v32i1.11776.

[3] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," *arXiv preprint*, arXiv:1906.08172, 2019. Available: https:// arxiv.org/abs/1906.08172.

[4] M. J. Hussein, J. Yusuf, A. S. Deb, L. Fong, and S. Naidu, "An Evaluation of Online Proctoring Tools," *Open Praxis*, vol. 12, no. 4, pp. 509–525, 2020. doi: 10.5944/openpraxis.12.4.1113.

[5] N. Dilini, A. Senaratne, T. Yasarathna, N. Warnajith, and L. Seneviratne, "Cheating Detection in Browser-based Online Exams through Eye Gaze Tracking," in *2021 6th International Conference on Information Technology Research (ICITR)*, Moratuwa, Sri Lanka, 2021.

[6] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, Jan. 2021. doi: 10.1109/TPAMI.2019.2929257.

[7] A. Abozaid and A. Atia, "Multi-Modal Online Exam Cheating Detec- tion," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Prague, Czech Republic, 2022, pp. 1–6. doi: 10.1109/ICECET55527.2022.9873527.

[8] Roboflow, "Ex-Guard-2 Dataset," [Online]. Available: https://universe. roboflow.com/mini-project-ppkno/ex-guard-2/dataset/5.

[9] Roboflow, "Cheating Detection Dataset," [Online]. Available: https://universe.roboflow.com/mahmoud-mohamed-phhz1/cheating-vfvwa/dataset/1.