Ensemble YOLO Framework for Enhanced Exam Monitoring In Offline Examination

Shravan S S [#1] Niruthya Vaani B[#2] Visva R[#3]

shravan.ss2021@vitstudent.ac.in niruthyavaani.b2021@vitstudent.ac.in visva.r2021@vitstdent.ac.in

School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT), Chennai,

Tamil Nadu, India.

***Abstract*—The increasing incidence of cheating in academic examinations threatens the credibility of educational institutions, highlighting the need for effective monitoring systems. This paper presents an automated offline exam cheating detection system that leverages advanced computer vision techniques to identify suspicious behaviors in real-time by analyzing CCTV footage from examination halls. The system integrates two primary models: YOLO-NAS (You Only Look Once - Neural Architecture Search) for detecting cheating behaviors, such as the use of unauthorized materials like mobile phones and papers, and YOLOv8 for analyzing student actions, classifying behaviors into categories such as daydreaming, normal, and stressed. The models were trained on meticulously labeled datasets, with data labeling and augmentation facilitated through tools like Labelbox. The system demonstrated high performance, achieving a mean Average Precision (mAP) of 74.9%, precision of 73.9%, and recall of 82.4%, effectively identifying cheating behaviors while minimizing false positives. By combining object detection with emotional analysis, this approach reduces reliance on human invigilators and offers a more automated, reliable method for ensuring academic integrity. This solution represents a significant advancement in the application of AI and computer vision in educational settings, ensuring fairer assessments and fostering a culture of honesty in academic examinations.**

Keywords: YOLO model ensemble, exam monitoring, multimodal fusion, cheating detection, machine learning, audio-visual analysis, physiological data, academic integrity, suspicious behavior detection, real-time monitoring

I . Introduction

In recent years, academic institutions worldwide have faced an alarming rise in exam cheating, a trend that has been exacerbated by the shift to online and hybrid learning environments following the COVID-19 pandemic. This increase in dishonest behavior poses a significant threat to the integrity of education, undermining both individual achievements and institutional credibility. As a result, the need for innovative technological solutions to detect and prevent cheating in exams has become more critical than ever. In online exam settings, Internet of Things (IoT) devices and AI-based systems, such as Smart Detectors, have been increasingly implemented to monitor students. These systems leverage biosecurity protocols, anomaly detection, and continuous behavioral analysis to ensure exam integrity, aligning with the broader objectives of Society 5.0, which aims to enhance societal well-being through interconnected, intelligent systems designed for various applications, including education [1].

However, the challenges associated with offline exams are unique, as they often lack the continuous data and network connectivity that make online monitoring possible. Traditional proctoring methods in offline exam settings frequently fail to detect subtler forms of cheating, such as passing notes, whispering, or attempting to glance at a neighbor's paper. This gap in monitoring systems calls for the development of adaptable, accurate, and non-invasive detection solutions. To address this, machine learning and computer vision techniques, including deep learning algorithms such as Convolutional Neural Networks (CNNs) and You Only Look Once (YOLO) models, are being explored to identify suspicious behaviors like eye movement, head orientation, and body gestures during exams. These advanced systems provide real-time, automated detection, significantly reducing the potential for human error or oversight in monitoring. By leveraging these technologies, offline exam environments can achieve enhanced security and fairness while safeguarding student privacy, offering a promising solution to the ongoing issue of exam cheating [12][13][18].

II . Related works

Literature on cheating detection in examinations underscores the role of artificial intelligence (AI) in enhancing exam integrity and identifying dishonest behaviors. Traditional proctoring, although effective in controlled, smaller settings, often falls short in large-scale assessments where the reliance on human invigilators introduces subjectivity and potential bias. AI-based systems aim to address these limitations, providing automated, real-time capabilities that improve both the efficiency and accuracy of detecting cheating behaviors. Research in this domain encompasses various AI techniques, including computer vision, machine learning, and behavioral analytics, which collectively contribute to more robust and fair exam monitoring systems [1, 2].

A significant approach within AI-based cheating detection systems is behavioral analysis, which focuses on identifying suspicious actions through computer vision and machine learning algorithms. Behavioral cues such as eye movement,

head orientation, and body posture are captured through continuous monitoring of students during exams. These features are then processed to distinguish ordinary behaviors from those indicative of potential cheating. Studies have found that integrating gaze-tracking technologies with machine learning algorithms enhances detection accuracy significantly [3]. For instance, Rachmadi et al. (2020) demonstrated that head-pose estimation combined with gaze-tracking could detect cheating with a 20% higher accuracy compared to gaze-tracking alone. These systems effectively reduce false positives and highlight the utility of multi-feature analysis in cheating detection [4].

Moreover, AI-driven motion detection and pose estimation are increasingly used to monitor body language and gestures, which are difficult to assess in real time using traditionalmethods. For instance, researchers have developed systems that utilize skeleton mapping and body-joint tracking to identify specific actions associated with cheating, such as frequent turning or looking away from the screen. A recent study by Kaur et al. (2021) implemented a convolutional neural network (CNN)-based motion detection model, which accurately identified these behaviors, outperforming traditional rule-based detection systems in terms of reliability [5]. This method of tracking body language helps in creating a non-intrusive monitoring system, where AI tools seamlessly track motion without excessive user intervention or privacy concerns [6].

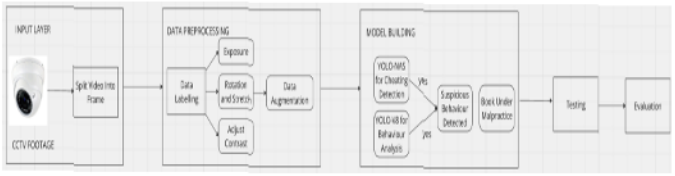
Audio analysis is another dimension within AI-based cheating detection. Audio-based systems analyze ambient sounds to detect whispers or unusual noises that may indicate communication with other individuals. By applying noise-canceling algorithms and speech recognition, AI can filter background noise to focus on identifying suspect conversations. This approach, highlighted in the work of Alotaibi and colleagues (2022), has proven to be effective in identifying non-verbal cues and suspicious sounds, thereby addressing an often- overlooked aspect of cheating detection [7]. Additionally, combining audio analysis with visual tracking systems has shown promising results in reducing detection error rates and increasing overall monitoring accuracy [8].

Recent developments in AI-based proctoring have also introduced the concept of multimodal fusion, where multiple data sources—such as video, audio, and biometric signals—are analyzed collectively to improve detection precision. This multimodal approach enables a more comprehensive understanding of students' actions, allowing the system to validate unusual behaviors detected by one mode against data from other sources. For instance, Wang et al. (2023) demonstrated that combining gaze tracking, pose estimation, and voice analysis led to a 30% increase in detection accuracy compared to single-modality systems [9]. This holistic method not only enhances reliability but also reduces false positives, contributing to a fairer and more accurate monitoring environment [10].

Lastly, ethical considerations and data privacy remain important topics within AI-based exam monitoring research. The intrusive nature of AI surveillance has raised concerns aboutstudents' privacy rights and the potential for over-surveillance. Researchers emphasize the need for ethical guidelines to ensure that AI systems respect students' privacy while effectively detecting cheating. Studies, such as those by Lee and Huang (2022), advocate for transparent AI frameworks that inform students about monitoring protocols and allow for appeals or reviews of detection results in cases of false positives. This ethical framework, combined with privacy-preserving AI techniques, aims to balance security with individual rights, ensuring that AI-based proctoring is both effective and just [11].

III. Proposed Methodology

The methodology for the offline exam cheating detection system is designed to systematically address the problem of detecting cheating and analyzing student behavior during examinations.



3.1 Data Collection

The methodology begins with the collection of video data from CCTV cameras strategically positioned in the examination hall. These cameras capture real-time footage of students during the exam, providing a rich source of information for behavior analysis. The recorded video is saved in a suitable format, ensuring that it can be efficiently processed in subsequent stages.

Tools and Libraries:

- Labelbox: A data labeling tool used to annotate video frames with labels for training the detection models, helping to identify suspicious behaviors.
- NumPy: A library for numerical computing in Python, used for handling arrays and performing mathematical operations during data processing.
- Pandas: A data manipulation and analysis library that provides data structures like DataFrames, facilitating data handling and analysis tasks.
- TensorFlow: An open-source deep learning framework that enables the training and deployment

of machine learning models, including YOLO models.

- YOLO-NAS: A variant of the YOLO model utilizing Neural Architecture Search to improve the detection of cheating behaviors, optimized for performance.
- YOLOv8: An upgraded version of the YOLO model designed for video segmentation and behavior analysis, capable of identifying emotional states such as daydreaming or stress.
- Matplotlib: A plotting library for creating static, animated, and interactive visualizations in Python, used for analyzing and displaying model performance metrics.
- scikit-learn: A machine learning library that provides tools for model evaluation and performance metrics, used for assessing the effectiveness of detection models.
- Flask: A micro web framework for building web applications, potentially used to create a dashboard for real-time monitoring and alerts.
- FFmpeg: A multimedia framework used for handling video, audio, and other multimedia files, which may assist in converting video formats or compressing video data.
- GPU/TPU (CUDA): Hardware accelerators (such as NVIDIA GPUs) used to speed up the training and inference processes for deep learning models.
- SQLite/MySQL: Database systems that may be used for storing annotated data and model results for later analysis and reporting.
- PyTorch: An open-source deep learning framework that may also be utilized for training models, depending on implementation preferences.

3.2 Data Preprocessing

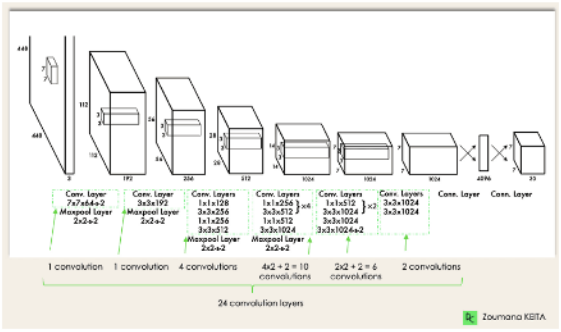
In the data preprocessing phase, the raw video footage is prepared for analysis through several steps:

- Frame Extraction: The recorded video is processed to extract individual frames using tools such as OpenCV. This step allows for detailed analysis of student behavior by examining each frame individually.
- Data Labeling: Each extracted frame is labeled using Labelbox, where specific behaviors are annotated. Labels indicate the presence of suspicious actions (e.g., looking at peers, using unauthorized materials) as well as neutral behaviors (e.g., normal, daydreaming, or stressed). This annotated dataset serves as the foundation for training the detection models.
- Data Augmentation: To improve the robustness of the models, data augmentation techniques are applied to the labeled frames. This includes rotations, flipping, and contrast adjustments to enhance image quality and increase dataset diversity.

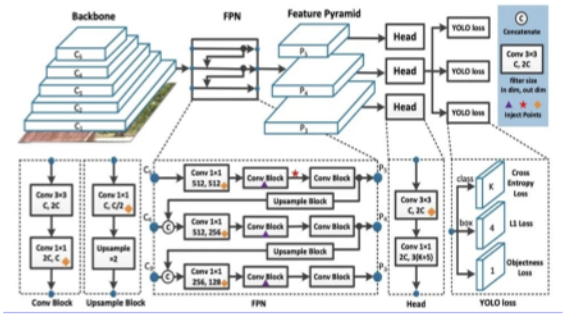
3.3 Model Training

The heart of the system lies in the model training phase, which involves developing two distinct models:

- Cheating Detection Model (YOLO-NAS): A YOLO-NAS (You Only Look Once - Neural Architecture Search) model is trained specifically for detecting cheating behaviors. This model classifies three main classes: cheating, paper, and non-cheating behaviors (such as using a phone). The training process involves feeding the labeled frames into the model, allowing it to learn the distinguishing features of each class.



- Behavior Analysis Model (YOLOv8): Simultaneously, a YOLOv8 model is employed for analyzing student behavior. This model categorizes behaviors into three classes: daydreaming, normal, and stressed. It is trained on a dataset of labeled frames, enabling it to accurately recognize emotional states and actions of students during the examination.



3.4 Testing and Evaluation

Following model training, a rigorous testing and evaluation phase is conducted:

- Testing Dataset: A separate dataset is utilized for testing, consisting of frames that the models have not encountered during training. This allows for an unbiased evaluation of the models' performance in detecting suspicious behaviors.
- Performance Metrics: The effectiveness of both models is assessed using various performance metrics, including precision, recall, F1-score, and accuracy. These metrics provide valuable insights into the models' capabilities in real-world applications and their reliability in flagging potential cheating incidents.

3.5 Integration and Real-Time Monitoring

The final phase involves the integration of both models into a comprehensive system capable of real-time monitoring during examinations:

- Real-Time Analysis: The system continuously analyzes incoming video feeds from the CCTV cameras. The YOLO-NAS model identifies any instances of cheating, while the YOLOv8 model evaluates student behavior to detect signs of stress or distraction.
- Alert Mechanism: Based on the combined outputs of both models, an alert system is triggered when suspicious activities are detected. The results are compiled and presented to exam invigilators, facilitating timely intervention and investigation.

3.6 Continuous Improvement

The methodology emphasizes the importance of continuous improvement. Post-deployment, the system is monitored for performance, and feedback from invigilators is used to enhance the models further. Regular updates and retraining sessions will be conducted to adapt to evolving cheating methods and ensure the system remains effective.

3.7 Implementation

The implementation of the offline exam cheating detection system follows a structured approach, ensuring that each component is effectively developed and integrated. The key steps are outlined below:

A. Environment Setup

- Install Python: Ensure the latest version of Python is installed on the development machine.
- Install Required Libraries: Use a package manager to install essential libraries such as:
    - OpenCV for video processing
    - TensorFlow or PyTorch for model training
    - Labelbox for data labeling
    - NumPy for numerical operations

B. Data Collection and Preparation

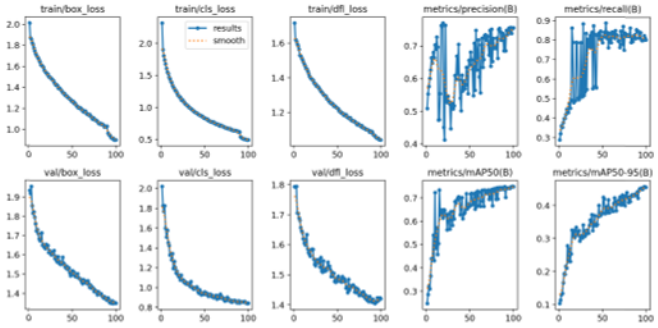
- Collect Video Data: Capture video footage from CCTV cameras placed in the examination hall.
- Extract Frames:
    - Input the video file.
    - Loop through each frame of the video.
    - Save each frame as an image file for analysis.
- Label Data:
    - Use Labelbox to annotate each frame.
    - Define labels for suspicious behaviors (e.g., cheating, paper, non-cheating, daydreaming, normal, stressed).
    - Export the labeled dataset for training.

C. Model Training

- Train YOLO-NAS for Cheating Detection:
    - Initialize the YOLO-NAS model with defined classes for detection (cheating, paper, non-cheating).
    - Train the model using the labeled dataset.
- Train YOLOv8 for Behavior Analysis:
    - Initialize the YOLOv8 model with defined classes for behavior analysis (daydreaming, normal, stressed).
    - Train the model using the labeled behavior dataset.

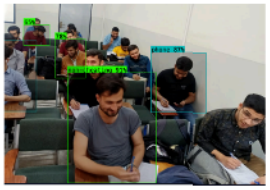
D. Testing and Evaluation

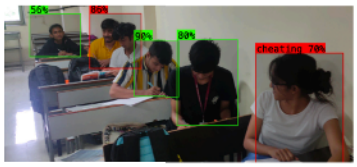
- Evaluate Models:
    - Use a separate testing dataset to assess the performance of both models.
    - Calculate metrics such as accuracy, precision, recall, and F1-score for each model to validate their effectiveness.



E. Real-Time Integration

- Set Up Real-Time Video Processing:
    - Input the live video feed from the CCTV cameras.
    - Continuously read frames from the video feed.
- Cheating Detection:
    - For each frame, apply the YOLO-NAS model to detect cheating behaviors.
- Behavior Analysis:
    - For the same frame, apply the YOLOv8 model to analyze student behavior.
- Process Results:
    - Check the results from both models for any flagged suspicious activities.
    - If cheating or suspicious behavior is detected, trigger an alert for invigilators.





F. Continuous Monitoring and Improvement

- Monitor System Performance:
  - Continuously collect data on the system's performance during exams.
  - Analyze feedback from users and exam outcomes to identify areas for improvement.
- Update Models:
  - Regularly retrain models with new data to improve accuracy and adapt to changing behaviors.

G. Documentation and User Training

- Create Documentation:
  - Write comprehensive documentation detailing system operations, troubleshooting steps, and guidelines for interpreting results.
- Conduct Training:
  - Organize training sessions for exam invigilators to ensure they are familiar with the system's operation and capabilities.

IV. Results and Discussions

The automated exam cheating detection system achieved strong performance, with a mean Average Precision (mAP) of 74.9%, a precision of 73.9%, and a recall rate of 82.4%. These metrics demonstrate the model's effectiveness in accurately identifying and classifying cheating behaviors, with a high likelihood of detecting true cheating instances and minimizing false positives. The mAP indicates solid overall performance, while precision and recall reflect the model's ability to balance detection accuracy and coverage, ensuring reliability in monitoring academic integrity.

Training results showed consistent improvements, with both training and validation losses declining, signaling effective learning and minimal overfitting. Notably, mAP50 and mAP50-95 showed significant progress, suggesting the model's robustness in detecting behaviors across diverse, unseen data. Although some fluctuations in training losses occurred, this was likely due to task complexity or optimization challenges. Further refinements, such as hyperparameter tuning and data augmentation, could improve the model's performance, particularly in enhancing the mAP50-95 score and minimizing minor training loss fluctuations

V. Conclusion and Future Works

This study presents an innovative approach to cheating detection in examination settings by utilizing YOLO-NAS for object detection and YOLO-VS for behavioral analysis, demonstrating an mAP of 74.9%, precision of 73.9%, and recall of 82.4%. The system effectively identifies suspicious behaviors, supporting examination integrity and reinforcing the literature advocating for automated monitoring in educational environments. By reducing the reliance on human invigilators, this technology streamlines exam processes and enhances fair assessment practices. However, challenges remain, including improving the mAP50-95 score and addressing fluctuations in training losses, emphasizing the need for meticulous data preparation and robust architecture. Future work will focus on expanding the dataset, refining hyperparameters, and exploring alternative model architectures. Additionally, integrating facial recognition for automatic student identification, saving video footage with bounding boxes around detected faces, and advancing video analysis capabilities to detect specific cheating patterns will further improve the system. This will allow for more accurate tracking of individual behaviors and ensure accountability with concrete evidence of suspicious activity. As part of future improvements, efforts will include refining facial recognition through expanded datasets and fine-tuning, along with generating comprehensive reports for each detected instance of suspicious behavior, which will be sent to the examination controller for evaluation and potential charges of malpractice. These findings underscore the potential of leveraging cutting-edge technology to uphold academic integrity, with ongoing research likely to yield even more effective solutions for exam monitoring.

VII. References

[1] T. Putra, S. Socrates, R. Fahlevi, I. Riziqul, M. Irvan, and W. Wahyudhie, "Smart Detector: Anti-Cheating Exam Detection Based on IoT and Biosecurity Management to Make a Cultural Indonesia Towards a Society 5.0 Era," Spectrum, 2022, https://doi.org/10.54482/spectrum.v1i02.169.

[2] M. A. Insha, T. S. Naidu, R. S. R., and A. R. Kamble, "Detection of Malpractice in Offline Examination Using Deep Learning," 2022, https://doi.org/10.53759/aist/978-9914-9946-1-2_29.

[3] F. Kamalov, H. Sulieman, and D. S. Calonge, "Machine learning based approach to exam cheating detection," PLOS ONE, 2021, https://doi.org/10.1371/JOURNAL.PONE.0254340.

[4] K. P. Kamble and V. R. Ghorpade, "Video Interpretation for Cost-Effective Remote Proctoring to Prevent Cheating," 2021, https://doi.org/10.1007/978-981-33-4073-2_25.

[5] G. Henry, "WeCheat: Algorithm for e-Learning Smart Cheating Detection Using Mean-Shift Clustering," Smart Innovation, Systems and

Technologies, 2022, https://doi.org/10.1007/978-981-19-3112-3_31.
[6] G. S. Devi, G. Suvarna, and S. Chandini, "Automated Video Surveillance System for Detection of Suspicious Activities during Academic Offline Examination," World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 2017.
[7] R. Bawarith and A. Abdullah, "E-exam Cheating Detection System," International Journal of Advanced Computer Science and Applications, 2017, https://doi.org/10.14569/IJACSA.2017.080425.
[8] X. Chen and Q. Huang, "Anti-cheating remote online examination method and system," 2013.
[9] R. Lee and B. Blaszczyk, "Proctoring System for Remote Users Based on Machine Learning Techniques," 2022, https://doi.org/10.1007/978-981-19-5221-0_3.
[10] E. M. Imah, R. D. I. Puspitasari, F. Q. Annisa, and H. A. Habib, "A Comparative Study of Deep Transfer Learning Algorithm for Cheating Detection in the Exam Based on Surveillance Camera Recording," 2023, https://doi.org/10.1145/3626641.3626931.
[11] W. M. Alsabhan, "Student Cheating Detection in Higher Education by Implementing Machine Learning and LSTM Techniques," Sensors, 2023, https://doi.org/10.3390/s23084149.
[12] R. Saravanan, "Automatic Cheating Detection In Exam Hall," 2023, https://doi.org/10.36227/techrxiv.24538150.v1.
[13] R. Saravanan, "Automatic Cheating Detection In Exam Hall," 2023, https://doi.org/10.36227/techrxiv.24538150.
[14] A. Arinaldi and I. M. Fanany, "Cheating video description based on sequences of gestures," 2017, https://doi.org/10.1109/ICOICT.2017.8074679.
[15] S. Sanz, M. Luzardo, C. García, and F. J. Abad, "Detecting Cheating Methods on Unproctored Internet Tests," Psicothema, 2020, https://doi.org/10.7334/PSICOTHEMA2020.86.
[16] H. Agh and H. Hassanpour, "Abnormal Behavior Detection in Electronic-Exam Videos Using BeatGAN," 2022, https://doi.org/10.1109/ICSPIS56952.2022.10043922.
[17] M. Asad, M. Abbas, A. Asim, H. Ahmed, M. A. Munazza, A. Sadaf, M. A. Haq, and M. A. Asif, "Suspicious Activity Detection During Physical Exams," Social Science Research Network, 2023, https://doi.org/10.2139/ssrn.4676389.
[18] H. Sokiya, S. Jamankar, and M. Sonkar, "Using Video Surveillance for Cheating detection in Exam Hall," International Journal For Multidisciplinary Research, 2024, https://doi.org/10.36948/ijfmr.2024.v06i02.19174.
[19] P. Thuong-Cang, P. Anh-Cang, and T. Ho-Dat, "Exam Cheating Detection Based on Action Recognition Using Vision Transformer," Communications in Computer and Information Science, 2023, https://doi.org/10.1007/978-981-99-7649-2_6.
[20] S. Monteiro, R. Bhate, L. Sharma, and P. Shaikh, "Proct-Xam – AI Based Proctoring," 2022, https://doi.org/10.1109/ASIANCON55314.2022.9908817.
[21] M. Asadullah and N. Shibli, "An automated technique for cheating detection," 2016, https://doi.org/10.1109/INTECH.2016.7845069.
[22] M. Asadullah and N. Shibli, "An automated technique for cheating detection," 2016.
[23] P. Gadhave, "Smart Exam Proctoring System," International Journal For Science Technology And Engineering, 2023, https://doi.org/10.22214/ijraset.2023.51358.
[24] R. M. Alairaji, I. A. Aljazaery, H. T. Salim, and A. H. M. Alaidi, "Automated Cheating Detection based on Video Surveillance in the Examination Classes," International Journal of Interactive Mobile Technologies, 2022, https://doi.org/10.3991/ijim.v16i08.30157.
[25] J. H. Park, "Touch screen based CBT machine against cheating on the exam," 2005.
[26] "Intelligent Remote Online Proctoring in Learning Management Systems," 2023, https://doi.org/10.1007/978-981-19-7447-2_21.
[27] S. Suma, "Malpractice Detection Using Machine Learning," International Journal of Innovative Research in Advanced Engineering, 2023, https://doi.org/10.26562/ijirae.2023.v1005.03.
[28] A. Ghosh, A. Nhavalore, N. Ramaswamy, T. Shetty, P. L., and P. Auradkar, "Remote Proctoring System Using Video, Audio and System Scenario Analysis," 2022, https://doi.org/10.1109/ccem57073.2022.00009.
[29] D. M. Genemo, "Suspicious activity recognition for monitoring cheating in exams," Proceedings of the Indian National Science Academy, 2022, https://doi.org/10.1007/s43538-022-00069-2.
[30] H. Meng and Y. Ma, "Machine Learning–Based Profiling in Test Cheating Detection," Educational Measurement: Issues and Practice, 2023, https://doi.org/10.1111/emip.12541.