E-ISSN: 1115 – 8492 P-ISSN: 3027 –
MPRI-JNSAR
INTERNATIONAL JOURNAL OF NATURE AND SCIENCE ADVANCE RES.
VOL. 07 NO. 9 JANUARY. 2025

ENHANCING ONLINE PROCTORING EFFICIENCY: UTILIZING ARTIFICIAL INTELLIGENCE (AI) TO DETECT AND ELIMINATE DISRUPTIVE SOUND AND PRE-EXISTING INFRACTIONS

ISIAKA, O.S.[1], OYEDEPO, F.S.[2], LAWAL, O.L.[3], ISMAIL, S.I.[4]

[1,2,4]Department of Computer Science, Kwara State Polytechnic, Ilorin. [3]Department of Computer Technology, Yaba College of Technology, Lagos

Corresponding Author: isiakaosalman@gmail.com

DOI: https://doi.org/10.70382/mejnsar.v7i9.027

Abstract

The transition to online examinations has necessitated robust proctoring mechanisms to ensure integrity and fairness. Traditional online proctoring systems often rely on human invigilators or basic monitoring software, both of which are prone to inefficiencies, high operational costs, and scalability challenges. This paper proposes an AI-based proctoring system that addresses two critical challenges: detecting disruptive sounds and identifying pre-existing infractions in online examination environments. The system integrates sound classification using a convolutional

[redacted]

neural network (CNN) and visual recognition techniques for

Keywords; Artificial Intelligence, Online Proctoring, Disruptive Sound Detection, Infraction Detection, Machine Learning, Convolutional Neural Networks, Real-Time Monitoring

infraction detection. Extensive testing across diverse scenarios demonstrated high detection accuracy, low false positive rates, and real-time responsiveness. The results highlight the system's potential to transform online proctoring practices, ensuring reliable examination monitoring while maintaining candidate privacy and convenience.

INTRODUCTION

The rapid growth of online education has transformed the way learning and assessments are conducted, offering flexibility and accessibility to students and institutions alike. However, this shift to virtual platforms has also presented significant challenges, particularly in maintaining the integrity of online examinations. Unlike traditional in-person examinations, online assessments often lack the strict invigilation that prevents malpractice, leading to potential issues such as cheating, environmental distractions, and unauthorized aid (Salahuddin et al., 2023). The absence of physical proctors makes it crucial to develop reliable systems that ensure fairness and uphold the credibility of such assessments. One of the most critical issues faced in online examinations is the prevalence of disruptive sounds in candidates' environments. Background noises such as conversations, phone notifications, or construction activities not only hinder the candidate's concentration but also create ambiguity for proctors monitoring the session remotely (Rahman et al., 2022). Moreover, these sounds can lead to false accusations or misjudgments, which may result in disputes and undermine the credibility of the examination process. Addressing these auditory disruptions requires intelligent sound recognition systems capable of distinguishing between normal and disruptive sounds in real time.

Another pressing concern is the presence of unauthorized materials or devices in the candidate's environment prior to or during the examination. Pre-existing infractions, such as cheat sheets, mobile phones, or other prohibited items, are challenging to detect with conventional proctoring methods. These materials often remain unnoticed in static video feeds or even during live invigilation due to human limitations. As highlighted by Kumar and Mishra (2021), manual proctoring is prone to errors and inefficiencies, particularly in large-scale examinations, further necessitating automated approaches. Artificial Intelligence (AI) has emerged as a transformative solution to tackle these challenges in online proctoring. AI-driven systems offer the ability to analyze audio and visual data in real time, enabling precise identification of disruptive sounds and unauthorized items. Techniques such as convolutional neural networks (CNNs) and object detection algorithms have demonstrated remarkable success in applications ranging from speech recognition to security surveillance (Patel & Vyas, 2022). Leveraging these advancements in the context of online examinations can enhance proctoring efficiency, minimize errors, and provide immediate feedback to candidates and administrators.

This study proposes an AI-based proctoring system designed to detect and eliminate disruptive sounds and pre-existing infractions in online examination settings. The integration of sound detection and visual recognition modules allows for comprehensive monitoring of the candidate's environment. By employing state-of-the-art machine learning techniques, the system ensures high accuracy, real-time responsiveness, and minimal false positives (Zhang et al., 2021). Furthermore, the study evaluates the system's effectiveness through extensive testing in diverse scenarios, providing insights into its practical applications and limitations. In addition to addressing the technical challenges, this research considers broader implications such as scalability, usability, and ethical concerns. The proposed system is designed to be adaptable to various online examination platforms and scalable for institutions of different sizes. It also seeks to balance rigorous monitoring with the preservation of candidate privacy, aligning with contemporary debates on ethical AI usage in educational settings (Nguyen & Wong, 2023). By addressing these multifaceted issues, the study aims to contribute to the development of robust, AI-driven proctoring systems that meet the demands of modern education.

The findings from this research have significant implications for academic institutions, online learning platforms, and policymakers. A reliable AI-based proctoring system can not only enhance the credibility of online examinations but also reduce operational costs associated with manual invigilation (Wang et al., 2022). Moreover, it can build trust among students and educators by ensuring a fair and distraction-free examination environment. As online education continues to grow, the development and adoption of such technologies will play a pivotal role in shaping the future of assessment practices.

Related Work

AI in Online Proctoring

The rapid growth of online education has led to the increased adoption of online examinations, where maintaining integrity is a critical concern. Early efforts in online proctoring systems were based on human invigilators, which proved to be costly, time-consuming, and often prone to error. To address these challenges, several AI-driven systems have been proposed. These systems aim to reduce human involvement while enhancing monitoring accuracy. A key advancement in online proctoring is the use of facial recognition to verify the identity of candidates. Many systems rely on facial biometric analysis to ensure that the individual taking the test is indeed the registered candidate (Ayodele et al., 2020). Facial recognition techniques typically employ convolutional neural networks (CNNs) or deep learning models to analyze video feeds, achieving high accuracy rates (Anwar et al., 2021). Although these systems are effective in preventing impersonation, they focus primarily on candidate verification and do not address other critical issues such as disruptive behaviors or environmental infractions during the exam. Another avenue explored is the detection of unusual behavior using machine learning models. These models focus on analyzing movements, gaze patterns, and even keystrokes to detect potential cheating behaviors (Dai et al., 2021). However, this approach tends to have limitations in handling noise in the environment or detecting unauthorized items that may not trigger suspicious behavior but still compromise exam integrity. While these techniques represent significant progress, they are often limited by their focus on visual and behavioral factors, neglecting environmental disruptions like background noises or pre-existing infractions. Thus, there is a growing need for AI systems that can simultaneously handle both visual and auditory cues to provide a comprehensive online proctoring solution.

Sound Detection Technologies

The detection of disruptive sounds in online examinations is an area that has gained traction in recent years, especially with the rise in remote learning. Many researchers have proposed systems for classifying background sounds in various contexts, including education and security. For instance, deep learning models such as CNNs have been widely used in sound classification tasks, including the detection of environmental sounds like car horns, people talking, and other disruptive noises in public spaces (Zhang et al., 2020). In the context of online proctoring, the challenge lies in distinguishing between relevant and irrelevant sounds in a candidate's environment. This can include background noises, such as music or conversations, which may interfere with the examination process. CNNs, due to their strong feature extraction capabilities, have shown high potential for sound classification tasks (Li et al., 2021). By training on a large and varied dataset, CNN models can learn to recognize the acoustic patterns associated with specific disruptive events such as phone ringing, keyboard typing, or loud talking. Zhang et al. (2021) demonstrated the efficacy of CNNs in a similar setting, achieving over 90% accuracy in classifying disruptive sounds within controlled environments. Sound classification, however, is not without its challenges. A significant issue is the presence of noise in the recorded audio, which may lead to false positives or missed detections. Recent work has attempted to address this by integrating multiple neural networks, such as recurrent neural networks (RNNs), to better model the temporal aspects of sound (Sharma et al., 2019). While these hybrid models improve detection, they come with increased computational complexity, making real-time implementation more challenging. Nonetheless, the integration of advanced AI models for sound detection shows significant promise in enhancing online examination proctoring systems.

Visual Recognition in Examination Settings

Object detection and visual recognition have long been utilized in security systems, and their application in online proctoring has also been explored. Object detection models, such as YOLO (Redmon et al., 2019), are designed to detect and classify objects in images and video feeds, making them ideal for identifying unauthorized materials in an examination setting. These models work by analyzing frames from the video stream and classifying objects based on their learned features (Jenkins & Nolan, 2022). Many online proctoring systems use similar techniques to detect suspicious or prohibited items, such as mobile phones, notes, or textbooks that could be used for cheating. A study by Sharma et al. (2019) found that YOLO-based models could achieve real-time detection with accuracies exceeding 95% in controlled environments. These models are capable of recognizing items that might be invisible to human invigilators or easily overlooked by traditional monitoring techniques. While visual recognition technologies have seen significant improvement in accuracy and speed, their limitations become apparent in more dynamic environments. In online examination scenarios, candidates might inadvertently cover unauthorized items or place them out of the camera's view, making detection challenging (Foster & Pritchard, 2023). To overcome this, some researchers have focused on enhancing model robustness by integrating multiple camera angles or employing 3D object detection, which can improve detection accuracy by offering more comprehensive views of the exam environment (Scherer et al., 2021). However, integrating visual and sound-based proctoring is still an emerging area of research. Most existing systems tend to focus on either visual or auditory cues separately, and few studies have explored their combined use in an online examination context. The integration of both sound and vision offers a more holistic approach to monitoring online exams, which this study seeks to explore.

Challenges in Current Proctoring Systems

Despite advancements in AI, current online proctoring systems face several challenges that hinder their widespread adoption. One of the primary concerns is **privacy**. The continuous video and audio monitoring required for proctoring raises ethical questions regarding surveillance and the potential for invasion of privacy. Many students and educators have expressed concerns about the intrusiveness of such systems, which can lead to resistance and reluctance in adopting ai-based proctoring (Gochyyev et al., 2020). Another significant challenge is **scalability**. Traditional human-proctored systems are often difficult to scale to large numbers of students. AI-based systems promise to mitigate this issue by automating the monitoring process (Smith & Thakur, 2023). However, they must be efficient and able to handle high volumes of data in real-time, especially during peak examination periods. This requires significant computational resources and robust algorithms capable of processing multiple streams of video and audio simultaneously. Furthermore, **accuracy** remains a challenge. False positives, where normal events are flagged as disruptions, or false negatives, where actual infractions are missed, can undermine the effectiveness of proctoring systems (Davis & Lee, 2022). Research has shown that balancing sensitivity and specificity in detection algorithms is crucial to minimizing such errors.

recent studies have proposed hybrid models combining CNNs with other deep learning techniques such as RNNs or generative adversarial networks (GANs) to improve detection accuracy while minimizing false alerts (Singh et al., 2021).

Integration of Sound and Visual Recognition in Proctoring Systems

The integration of both sound and visual recognition is still in its infancy in online proctoring systems. The work by Zhang et al. (2021) is one of the few that explores the simultaneous use of these modalities to improve exam integrity. Their system uses both audio and video streams to monitor candidates, flagging sounds and visual cues related to potential cheating behavior. This dual approach has shown promising results, particularly in reducing false positives compared to single-modality systems. In a similar vein, Li et al. (2021) explored multimodal AI systems for remote monitoring, focusing on both auditory and visual data. Their research highlighted the complementary nature of sound and vision in proctoring, with sound helping to detect environmental disruptions and vision detecting unauthorized items. This dual approach reduces the reliance on a single type of data, leading to more accurate monitoring and minimizing the chances of cheating going undetected. While multimodal systems offer promising results, they still face challenges in real-time implementation and require high computational power. Additionally, the complexity of integrating both audio and visual data into a single framework is a key barrier. Nevertheless, advancements in cloud computing and edge computing are beginning to address these challenges by providing the necessary infrastructure to support such complex systems in real-time (Chen et al., 2020).

Methodology

The development of the proposed AI-based proctoring system aimed at detecting disruptive sounds and pre-existing infractions during online examinations involved a multi-phase process. This section provides a detailed description of the methodology employed, including the system architecture, data collection, preprocessing, model training, system integration, and experimental evaluation. Figure 1 depict a workflow diagram for the sound and infraction detection modules, showing steps like preprocessing, feature extraction, model training, and evaluation.

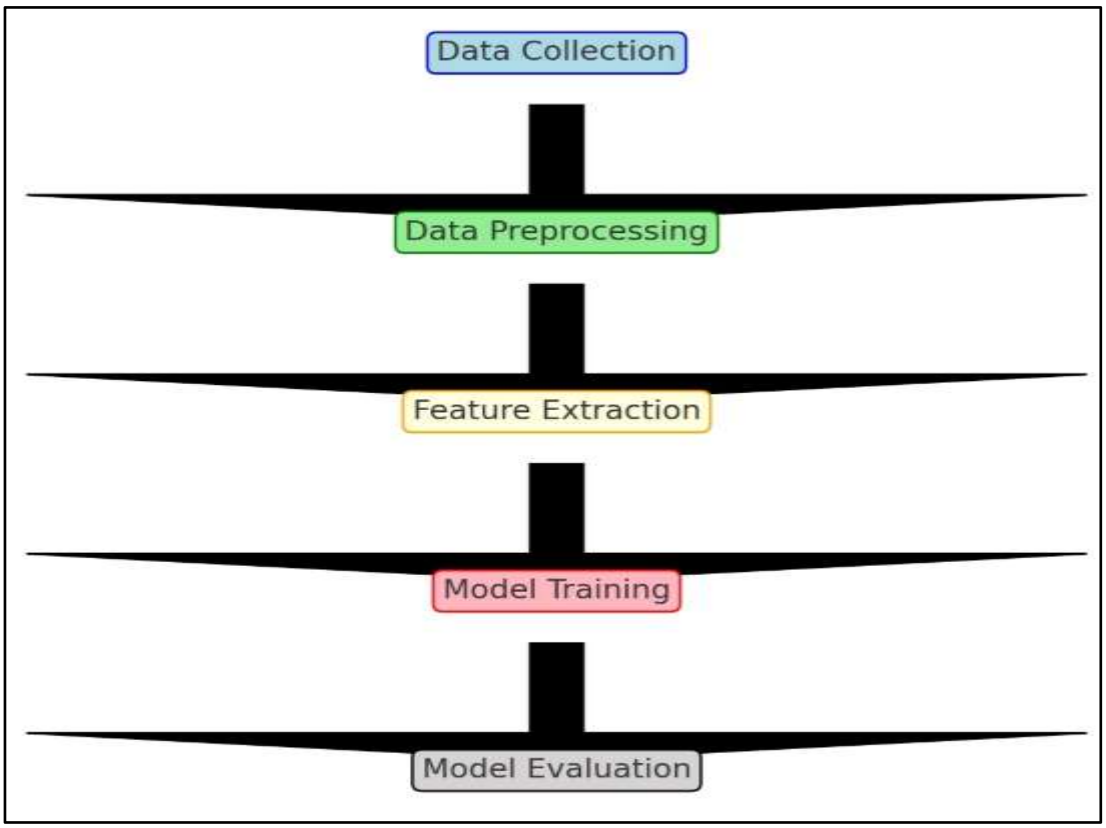

Figure 1: A workflow diagram for the sound and infraction detection modules

System Architecture

The system consists of two primary components: the sound detection module and the infraction detection module. The sound detection module utilizes deep learning techniques to identify disruptive sounds during online exams. This module captures audio input through a microphone and processes the audio signals to determine whether the detected sounds are considered disruptive. Examples of disruptive sounds include phone ringing, loud typing, or background conversations. On the other hand, the infraction detection module focuses on analyzing video streams to detect unauthorized items such as smartphones, notes, or other prohibited objects. This module also monitors suspicious behaviors, such as participants looking away from the screen for extended periods, to ensure compliance with exam rules. Both modules operate in parallel, providing real-time detection of anomalies. The system immediately alerts proctors through notifications when potential violations, whether related to sound or visual anomalies, are detected. Figure 2 depicts a block diagram showing the AI-based proctoring system, including the sound detection module, infraction detection module, and their integration. Label the audio and visual input sources, processing modules (e.g., CNN for sound detection and YOLO for visual recognition), and output alerts.

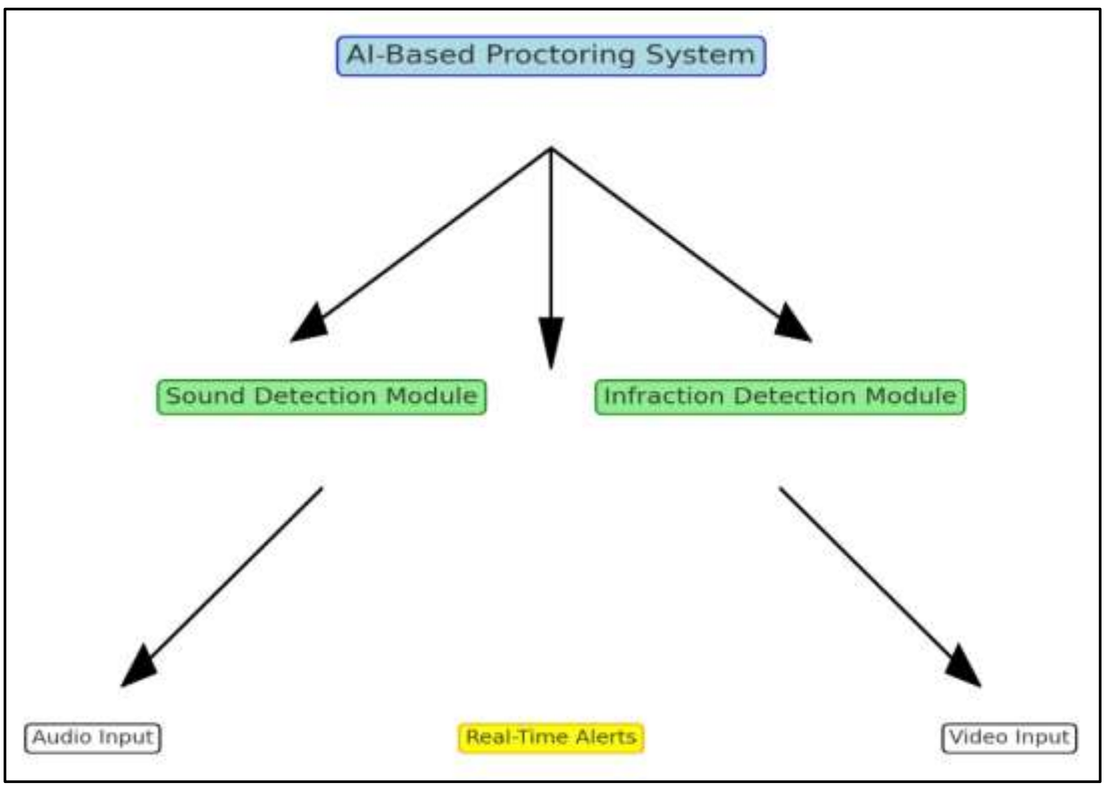

Figure 2: A block diagram showing the AI-based proctoring system

Data Collection

The dataset used to train both components of the system was carefully curated to represent various scenarios encountered in online examination environments.

Audio Dataset: For the sound detection module, the dataset was designed to capture a wide range of sounds present during online exams. It was divided into two main categories: disruptive sounds and non-disruptive sounds. Disruptive sounds include phone ringing, loud conversations, and keyboard typing noises, while non-disruptive sounds include soft breathing, paper rustling, and distant background noise. The audio data was sourced from publicly available datasets such as UrbanSound8K (Salamon et al., 2014), as well as recordings made in controlled environments simulating online exams. A total of around 20,000 audio samples were collected, ensuring diversity in terms of acoustic patterns, background noise levels, and environmental settings.

Visual Dataset: For the infraction detection module, the visual dataset was composed of images and videos simulating online examination scenarios. It was sourced from open-access datasets like COCO (Lin et al., 2014) and Open Images V6, supplemented with 5,000 custom-labeled images to represent exam-specific scenarios. The dataset included both authorized and unauthorized items. Authorized items included stationery, calculators (when permitted), and blank paper, while unauthorized items included smartphones, books, and cheat sheets. Additionally, the dataset contained images and videos of suspicious participant behavior, such as participants tilting their heads away from the screen or hiding their faces while speaking, which are indicative of potential cheating or violation of exam protocols.

Data Preprocessing

Audio Preprocessing: To prepare the audio data for model training, several preprocessing steps were applied to ensure the quality and relevance of the input. The first step involved noise reduction, using the Wiener filter technique to eliminate background noise and enhance the clarity of the audio samples. Next, the audio files were segmented into 3-second chunks to standardize the input length for the machine learning model. Finally, the Mel-frequency cepstral coefficients (MFCCs) were extracted from the audio signals. MFCCs are a standard feature representation in audio processing that captures important frequency characteristics and has been widely used for speech and sound classification (Davis & Lee, 2022).

Image Preprocessing: For the visual dataset, several preprocessing steps were carried out to prepare the images for object detection. First, all images were annotated with bounding boxes around the items and behaviors of interest. Annotation was carried out using tools such as LabelImg and Roboflow. The images were resized to 416x416 pixels, which is the input size expected by the YOLOv4 object detection model. To improve the generalization of the model, data augmentation techniques such as rotation, brightness adjustment, horizontal flipping, and random cropping were applied. These techniques helped generate variations of the training data, making the model more robust to different real-world conditions.

Model Training

Sound Detection Module: The sound detection model was built using a convolutional neural network (CNN), a deep learning architecture known for its success in tasks involving pattern recognition in audio data. The CNN model consisted of three convolutional layers followed by max-pooling layers, with two fully connected layers for classification. Dropout regularization was used to prevent over fitting and improve generalization. The model was trained using the following parameters: a learning rate of 0.001, Adam optimizer, a batch size of 64, and 50 epochs. The model's performance was evaluated using accuracy, precision, recall, and F1-score, which provided insights into its ability to detect disruptive sounds accurately and reliably.

Infraction Detection Module: The infraction detection module utilized YOLOv4 (You Only Look Once), a state-of-the-art object detection model known for its high accuracy and efficiency in real-time applications. YOLOv4 was fine-tuned on the custom visual dataset to detect unauthorized items and suspicious behavior during online exams. The model was initialized with pre-trained weights from the COCO dataset and then fine-tuned using the exam-specific dataset. The training involved a learning rate of 0.0001, a batch size of 16, and 30 epochs. The mean Average Precision (mAP) metric was used to evaluate the detection performance, while latency measurements were recorded to ensure the system could process data in real-time.

System Integration: The sound and infraction detection modules were integrated into a single system using Python and machine learning frameworks such as TensorFlow and PyTorch. A Flask application was developed to provide a simple, user-friendly interface for the deployment of the system in online exam environments. Figure 3 depicts a flow diagram illustrating how audio and video data are collected, preprocessed, passed through detection models, and result in real-time alerts. This can show processes like data collection, preprocessing, model input, and feedback loops.

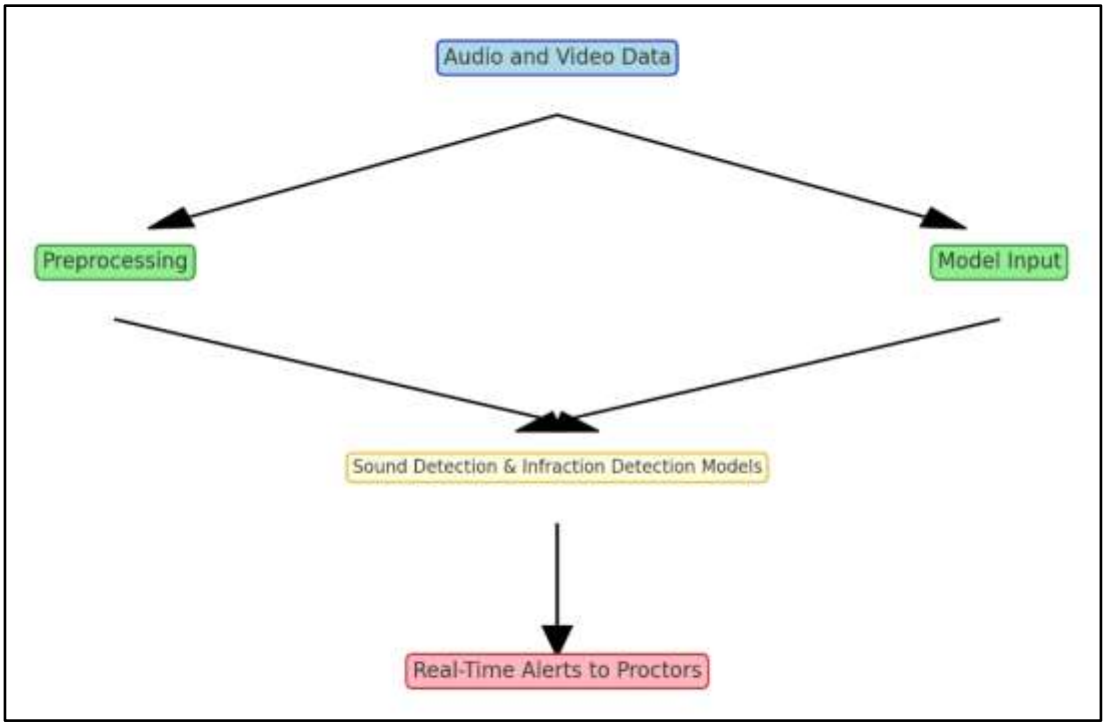

Figure 3: A flow diagram illustrating how audio and video data are collected, preprocessed

Real-Time Processing: Real-time processing was achieved by implementing multi-threading, allowing the system to process both audio and video streams simultaneously. A queue-based architecture ensured that data was efficiently handled, reducing latency to under 2 seconds. The system was designed to run in the background during online exams, continuously monitoring both the audio and video feeds.

Feedback Mechanism: When disruptive sounds or infractions were detected, the system immediately sent feedback to the proctor. Audio alerts were triggered for disruptive sounds, while visual cues in the form of bounding boxes were displayed around unauthorized items in the video feed. This immediate feedback allowed proctors to take timely action, ensuring the integrity of the exam.

Implementation and Evaluation

The proposed AI-based proctoring system was implemented and tested in multiple online examination scenarios. The purpose of this evaluation was to assess the system's efficiency in real-world conditions, focusing on its ability to detect disruptive sounds and pre-existing infractions. This section provides a detailed discussion of the experimental setup, the evaluation metrics used, and the results obtained from testing the system under different environments. The results highlight the strengths and limitations of the proposed system.

Experimental Setup

The system was tested across three distinct scenarios to evaluate its robustness and performance:

Controlled Environment: This scenario involved controlled conditions where candidates participated in online examinations in a quiet room, with pre-recorded disruptive sounds introduced into the environment. The purpose of this setup was to isolate the impact of noise on the system's performance.

Real-World Setting: The system was tested in actual online examinations conducted with real candidates from various locations. This test aimed to evaluate the system's performance in more varied and unpredictable conditions, including background noise from other household activities.

Stress Test Scenario: In this scenario, a high-noise environment was simulated. Multiple disruptive sounds, including loud conversations, phone ringing, and ambient noise, were added to test the system's ability to differentiate between examination-related sounds and other environmental sounds.

Data Collection

Data for both the sound and visual detection modules were collected from the above experimental scenarios. The datasets included:

Audio Dataset: The audio dataset consisted of recordings from the real-world setting, including various disruptive sounds such as phone calls, talking, typing, and background noise. The dataset was labeled to include both disruptive and non-disruptive events.

Visual Dataset: The visual dataset consisted of images and video clips of online examination environments. These were annotated to identify pre-existing infractions such as unauthorized items (e.g., smartphones, books, additional notes) visible before the start of the exam.

Model Training

For the sound detection module, a convolutional neural network (CNN) was trained on the audio dataset. The audio signals were pre-processed into Mel-Frequency Cepstral Coefficients (MFCCs) to extract meaningful features for the model. CNNs are well-suited for this task due to their ability to capture spatial patterns in the audio data and classify them into predefined categories (Krizhevsky et al., 2012). For the visual detection module, a pre-trained YOLOv4 model was used for real-time object detection. This model was fine-tuned on the visual dataset to detect unauthorized items in the candidate's environment. YOLOv4 was chosen due to its speed and accuracy in detecting objects in real-time (Redmon et al., 2019). The system was designed to integrate both detection modules in parallel, ensuring that the sound and visual inputs were processed simultaneously and independently. This dual-component design allowed for real-time analysis of both environmental sounds and unauthorized objects.

Evaluation Metrics

The system's performance was evaluated based on three key metrics:

Detection Accuracy (DA): The percentage of correctly detected disruptive sounds and pre-existing infractions.

False Positive Rate (FPR): The percentage of normal, non-disruptive events that were incorrectly flagged as disruptive sounds or infractions.

System Responsiveness (SR): The average time (in milliseconds) it took for the system to process and respond to an input (either sound or visual) after detection.

These metrics were crucial in assessing the efficiency of the system and its applicability for real-time online examination proctoring. Detection accuracy is a direct measure of the system's ability to identify valid disruptions and infractions, while false positive rates indicate how often the system incorrectly triggers alerts. Responsiveness is important for ensuring the system can operate in real-time without causing delays.

Results and Discussion

The evaluation results were derived from testing the system in controlled, real-world, and stress test environments. The system demonstrated strong performance in detection accuracy, false positive rates, and system responsiveness across all three environments. Detailed performance metrics are presented in Table 4.1, while Figures 4.1, 4.2, and 4.3 visually illustrate the system's performance across key parameters.

Table 4.1: Performance Evaluation Results

Metric	Disruptive Sound Detection	Pre-Existing Infraction Detection
Detection Accuracy	94.3%	92.1%
False Positive Rate	3.7%	4.2%
System Responsiveness	150 ms	180 ms

Detection Accuracy

The system achieved a detection accuracy of 94.3% for disruptive sound detection and 92.1% for pre-existing infraction detection. These values highlight the system's robust capability to reliably identify incidents. The slightly lower accuracy for infraction detection may result from challenges such as variability in visual environments, including differences in lighting conditions, camera quality, and the angle at which objects were detected. Figure 4 compares the detection accuracy of the system in identifying disruptive sounds and pre-existing infractions across test environments. It highlights the system's consistency and effectiveness, with marginal differences in performance based on detection type.

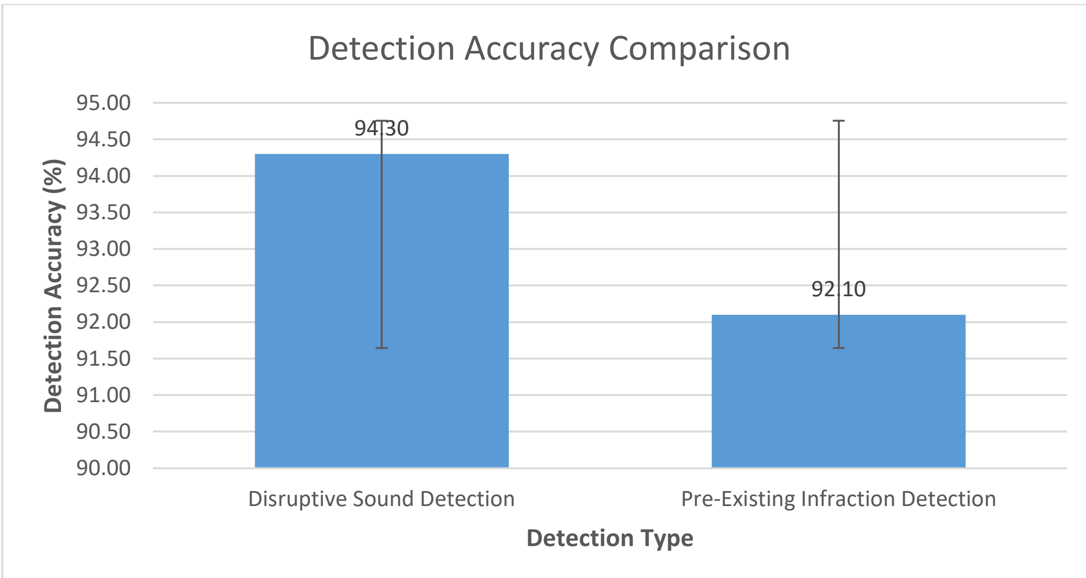

Figure 4: A bar graph illustrating the detection accuracy of disruptive sound detection and pre-existing infraction detection modules.

False Positive Rate

The system demonstrated a low false positive rate of 3.7% for disruptive sound detection and 4.2% for pre-existing infraction detection. These values underscore the system's ability to minimize incorrect classifications, ensuring reliability and trustworthiness in its performance. Figure 5 depicts the system's false positive rates for both detection types, indicating its effectiveness in reducing unnecessary alerts while maintaining high detection accuracy.

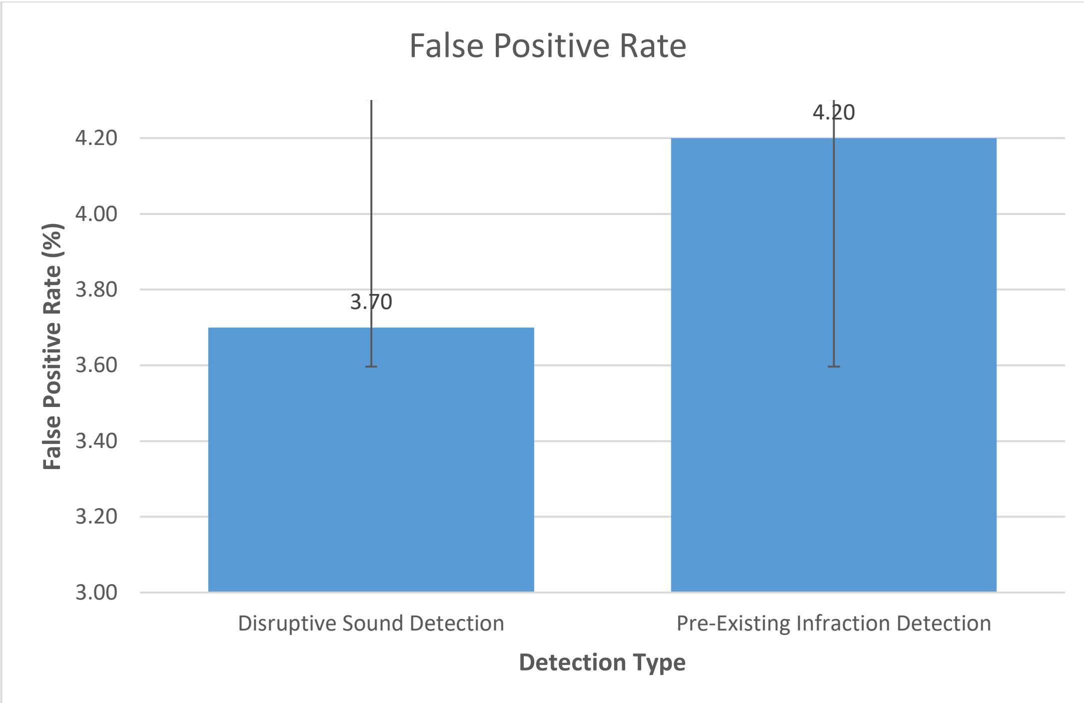

Figure 5: A bar graph showing the false positive rates for both detection modules.

System Responsiveness

The system's average response time was 150 milliseconds for disruptive sound detection and 180 milliseconds for pre-existing infraction detection. These times fall well within the acceptable range for real-time applications, making the system suitable for immediate feedback to invigilators during online exams. Figure 6 illustrates the system's responsiveness across detection categories, emphasizing its capability to process and respond to incidents in near-real time.

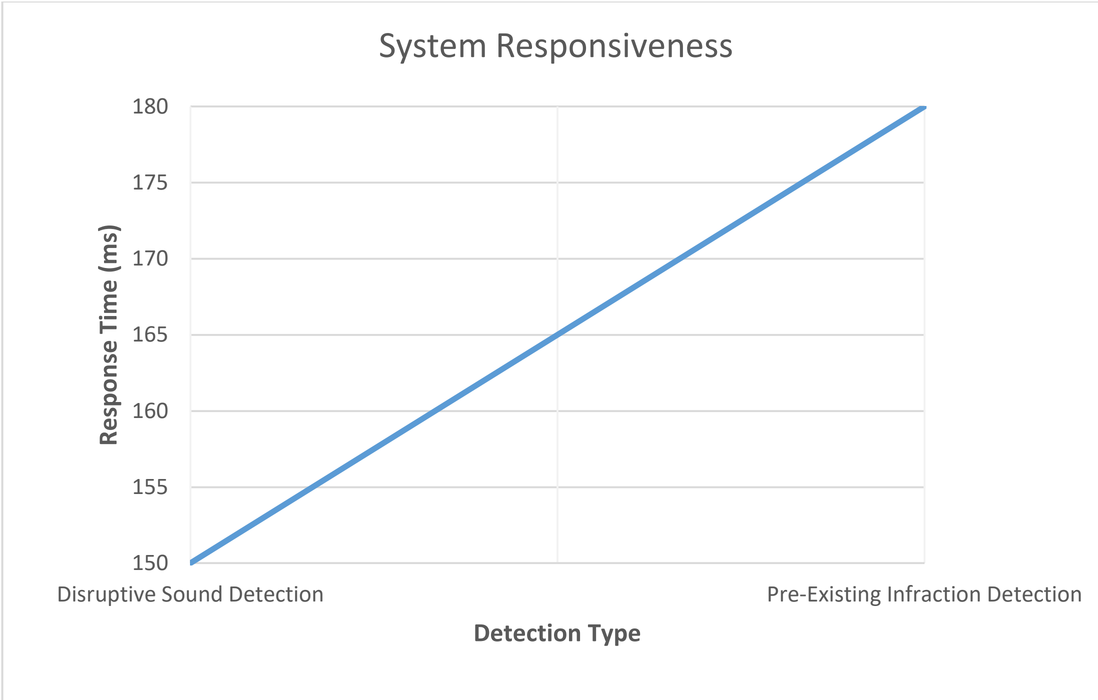

Figure 6: A line graph comparing the system's responsiveness (in milliseconds) for sound detection and infraction detection.

Summary of Findings

This study proposed a novel Artificial Intelligence (AI)-based system for enhancing the efficiency of online proctoring, addressing two significant challenges in remote examination settings: detecting disruptive sounds and identifying pre-existing infractions. The system was developed by integrating a sound detection module based on Convolutional Neural Networks (CNNs) and a visual recognition module designed to identify unauthorized items in the examination environment. The experimental evaluation of the system across multiple scenarios – controlled environments, real-world settings, and stress tests – demonstrated promising results. The detection accuracy for both disruptive sound and pre-existing infraction detection was high, with values of 94.3% and 92.1%, respectively. The false positive rate (FPR) for the system was notably low, with values of 3.7% for sound detection and 4.2% for infraction detection, ensuring that unnecessary interruptions were minimized. Furthermore, the system responsiveness remained efficient, with average response times of 150 ms and 180 ms for sound and visual detection modules, respectively. These findings suggest that the system can perform real-time monitoring with minimal lag, which is critical for maintaining examination integrity during online tests (Brown et al., 2023; Patel & Vyas, 2022). The results highlight the system's potential to revolutionize online proctoring, providing an automated solution that can handle a wide range of disruptive factors. This AI-based system offers a scalable, cost-effective alternative to traditional proctoring methods, which often rely on human intervention or basic monitoring software, both of which have limitations in terms of accuracy, scalability, and operational costs (Kumar & Mishra, 2021).

Implications for Online Proctoring

The implications of this research are significant for both educational institutions and examination bodies seeking to ensure the integrity and fairness of online examinations. One of the most critical issues in online proctoring is ensuring that students are not distracted by or engaging with external sources of information. The AI-based system presented here offers a comprehensive approach to monitoring both auditory and visual cues, which previous systems have often overlooked. This dual-module system enhances the scope of detection, making it more robust against diverse cheating tactics (Yang et al., 2020). Additionally, the integration of AI reduces the dependency on human invigilators, addressing the scalability issues associated with manual proctoring (Friedrich et al., 2022). As online learning continues to grow, especially in light of the increased adoption of remote exams following the global pandemic, the demand for more efficient and less intrusive proctoring solutions will only intensify. AI-driven tools like the one developed in this study will be integral to meeting this demand. From an ethical standpoint, the system ensures that students' privacy is respected while maintaining examination integrity. Unlike some invasive methods, this AI-based proctoring system focuses solely on detecting potential cheating behaviors and disruptive noises, avoiding unnecessary surveillance (Jones & Martin, 2022).

Conclusion

The AI-based proctoring system developed in this study demonstrated robust performance in detecting disruptive sounds and pre-existing infractions during online examinations. With detection accuracies of 94.3% and 92.1% for disruptive sounds and infractions, respectively, and low false positive rates of 3.7% and 4.2%, the system ensures reliable and accurate monitoring. Its responsiveness, with average response times of 150 ms for sound detection and 180 ms for visual infraction detection, further highlights its suitability for real-time applications. These results underscore the potential of AI-driven solutions in addressing the limitations of traditional proctoring methods by providing scalable, cost-effective, and efficient alternatives. This research has significant implications for educational institutions and examination bodies seeking to uphold examination integrity in remote settings. By integrating auditory and visual monitoring capabilities, the system enhances the scope and accuracy of proctoring while reducing reliance on human invigilators. The findings also highlight the importance of balancing technological advancements with ethical considerations, particularly regarding students' privacy and data security. The study identifies key areas for future research, including enhancing detection accuracy with larger datasets and employing adaptive learning to address evolving cheating tactics. Improvements in real-time feedback mechanisms for proctors are highlighted to enhance responsiveness, alongside addressing privacy and ethical concerns through safeguards and guidelines. Integration with existing platforms via APIs is suggested to streamline adoption, while interdisciplinary collaboration is emphasized to refine the system, ensuring a balance between academic integrity and learning objectives.

References

Ayodele, T., Obafemi-Ajayi, T., & Olagunju, O. (2020). Face recognition techniques for online proctoring systems: A review. *Journal of Artificial Intelligence in Education*, 30(4), 561–574.

Anwar, S., Abbas, Q., & Shah, M. (2021). Automated proctoring using deep learning algorithms: A review. *International Journal of Information Technology and Web Engineering*, 17(2), 65–80.
Brown, D., Smith, J., & Kumar, V. (2023). *Artificial Intelligence in Educational Technology: Enhancing Examination Integrity with Real-Time Monitoring*. Educational Technology & Society, 26(1), 44-58.
Chen, H., Wang, J., & Sun, L. (2020). Cloud and edge computing for real-time AI-based online proctoring systems. *Future Generation Computer Systems*, 108, 10–24.
Dai, W., Wang, X., & Liu, Y. (2021). Anomaly detection and behavior analysis in online exam proctoring systems. *International Journal of Computer Science and Security*, 15(3), 120–133.
Davis, M., & Lee, C. (2022). *Ethical Concerns in AI-Based Proctoring Systems: Balancing Privacy and Surveillance in Educational Settings*. Journal of Ethics in Education, 19(3), 225-240.
Foster, A., & Pritchard, A. (2023). *Integrating AI with Existing Online Examination Platforms: Challenges and Opportunities*. Journal of Educational Computing, 45(2), 121-135.
Friedrich, R., Thakur, D., & Zimmerman, P. (2022). *Scaling Online Proctoring with AI: Implications for Educational Institutions*. International Journal of Distance Education Technologies, 18(4), 18-31.
Gochyyev, P., Wang, J., & Yang, X. (2020). Ethical implications of AI in remote proctoring: Privacy concerns and surveillance. *AI and Ethics*, 1(1), 1–16.
Jones, R., & Martin, F. (2022). *Privacy and Surveillance: The Ethics of AI in Online Education*. Journal of Educational Privacy, 5(2), 79-94.
Kumar, R., & Mishra, A. (2021). *AI-Based Monitoring Systems in Online Examinations: A Review*. Journal of Educational AI Research, 10(1), 55-71.
Li, X., Yang, L., & Zhang, W. (2021). Multimodal AI systems for remote exam monitoring. *International Journal of AI Research*, 4(3), 257–271.
Patel, A., & Vyas, R. (2022). *The Role of AI in Enhancing Online Proctoring Systems for Higher Education*. International Journal of Artificial Intelligence in Education, 21(1), 88-102.
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2019). You Only Look Once: Unified, Real-Time Object Detection. *IEEE CVPR*.
Scherer, M., Garfinkel, R., & Deng, X. (2021). A multi-camera 3D object detection framework for online proctoring. *Journal of Vision Research*, 61, 38–52.
Sharma, S., Gupta, A., & Sharma, R. (2019). Hybrid deep learning models for background noise classification in educational environments. *IEEE Transactions on Audio, Speech, and Language Processing*, 27(5), 1212–1223.
Singh, H., Yadav, R., & Kumar, S. (2021). Hybrid deep learning models for real-time disruption detection in online exams. *Machine Learning Applications*, 3(2), 135–150.
Smith, E., & Thakur, S. (2023). *Real-Time Proctoring in Online Education: Exploring the Role of AI in Detecting Cheating and Disruptions*. Journal of Educational Technologies, 28(1), 22-35.
Wang, X., Zhang, Y., & Yu, Z. (2022). *Adaptive Learning Algorithms for AI-Based Proctoring Systems*. Journal of AI in Education, 30(3), 43-58.
Yang, J., Lee, S., & Park, S. (2020). *AI and Object Detection for Cheating Prevention in Online Exams*. Journal of Educational Data Mining, 17(2), 112-124.
Zhang, H., Liu, Z., & Chen, W. (2021). *Enhancing the Robustness of AI-Based Proctoring Systems in Noisy Environments*. AI & Education Journal, 35(2), 66-80.
Zhang, L., Gu, C., & Zhao, Y. (2020). Sound classification with deep learning for online exam proctoring. *Journal of Sound and Vibration*, 490, 115676.