J. Electrical Systems 20-3 (2024): 7375-7383

[1]Sangeeta Lamba
[2]Dr. Neelam Sharma

Deep Learning-Based Multimodal Cheating Detection in Online Proctored Exams

Abstract: - The rapid adoption of online learning has emphasized the need for reliable and scalable methods to ensure academic integrity during remote assessments. This study presents a sophisticated AI-powered e-cheating detection system that integrates state-of-the-art technologies such as YOLO-based face detection, Silero VAD for audio analysis, L2CS-Net for gaze tracking, and SixDRepNet for head pose estimation. Utilizing a hybrid CNN-BiLSTM architecture, the system processes webcam and audio data in real time to detect behaviors indicative of cheating, including unauthorized individuals, diverted attention, and off-screen audio communication. Key contributions include a modular framework for preprocessing sequential data, robust feature extraction methods, and a scalable architecture designed for high-dimensional behavioral analysis. The model achieves an accuracy of 87.5%, an F-score of 0.8762, and an AUC of 0.8795, highlighting its effectiveness in identifying cheating instances while minimizing false positives and negatives. Precision-recall and ROC curves validate the model's performance under varying operational thresholds, demonstrating its adaptability to real-world applications. This research underscores the potential of AI-driven solutions to mitigate the challenges of remote assessments, offering a balance between robust security and ethical considerations. The modular design ensures scalability and flexibility, making it suitable for deployment in diverse educational and certification environments. By advancing the capabilities of online exam proctoring systems, this study contributes to fostering trust and fairness in digital learning landscapes. Future work will focus on enhancing multimodal analysis and integrating privacy-preserving techniques.

Keywords: *Academic integrity, CNN-BiLSTM, YOLO, L2CS-Net, Behavioral analysis.*

I. INTRODUCTION

The rise of online education and remote testing has significantly increased the demand for innovative technologies to uphold academic integrity while ensuring a seamless and fair assessment experience. This study presents an Advanced Online Exam Proctoring System that utilizes state-of-the-art artificial intelligence to monitor and evaluate test-taker behavior during examinations. As of 2021, over 35% of U.S. postsecondary students were enrolled in at least one online course, compared to 15.6% in 2012 reported by National Center for Education Statistics [1]. This increase reflects the rising adoption of online learning environments, driven in part by technological advancements and the flexibility offered by distance education programs. The global online exam proctoring market is anticipated to grow from approximately USD 820.98 million in 2023 to USD 2,832.14 million by 2031, registering a compound annual growth rate (CAGR) of 16.7% during the forecast period (2023–2031) [2]. This expansion is primarily driven by the increasing adoption of e-learning solutions and the rising demand for effective online assessment systems. These trends reflect the growing emphasis on secure and scalable technologies for remote evaluations.

Virtual exams are online assessments that allow students to take tests remotely. They are designed to evaluate students' knowledge and application of course materials. To maintain exam integrity, many institutions use AI-powered proctoring systems, which monitor student behavior through technologies like facial recognition, screen tracking, and gaze detection. The popularity of virtual exams has surged, especially with the rise of online education and the impact of the COVID-19 pandemic. While they offer flexibility, these exams also raise concerns about academic integrity and cheating. AI-based proctoring systems help address these issues by providing real-time monitoring and identifying suspicious activities. However, the use of such technologies also brings privacy and security concerns, prompting discussions about how to balance robust security with ethical considerations. Key features include facial recognition, gaze tracking, object detection, and audio analysis, which work together to detect and mitigate cheating activities in real-time, such as unauthorized individuals, diverted attention, or the presence of prohibited items.

[1] *Corresponding author: Research Scholar, Department of Computer science, Banasthali Vidyapith, Jaipur, Rajasthan, India, 302044

[2] Associate Professor, Department of Computer science, Banasthali Vidyapith, Jaipur, Rajasthan, India, 302044

A distinguishing feature of this system is the application of the You Only Look Once (YOLO) algorithm, renowned for its speed and accuracy in object detection, which enhances the reliability of real-time monitoring. Additionally, the system is designed to balance robust security measures with ethical considerations, emphasizing the protection of user privacy and obtaining informed consent. Performance evaluations demonstrate that the system delivers high accuracy and scalability, making it suitable for large-scale use by educational institutions and certification providers. By reducing the reliance on human invigilation, the system minimizes biases and operational costs while promoting fairness and transparency. This research contributes to the ongoing development of secure, unbiased, and efficient online examination systems, paving the way for advancements in AI-driven educational technologies. The system demonstrated the capability to process webcam data, detect critical user behaviors, and provide reliable and scalable solutions for real-time applications like proctoring or behavioral analysis.

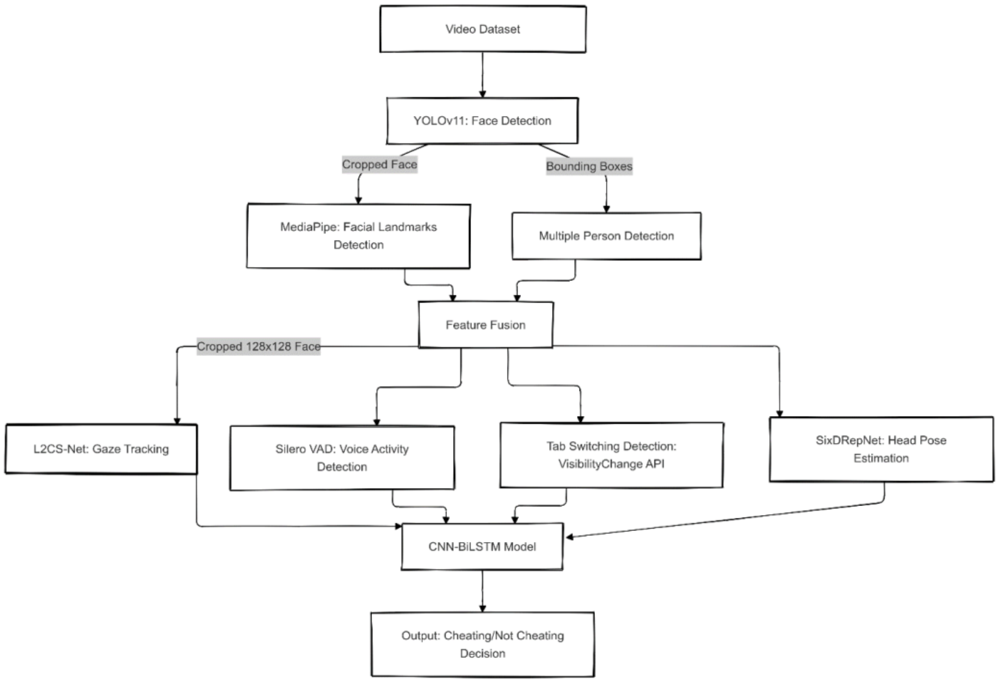

Fig. 1 Proposed Framework for e-exam proctoring environment

The structure of this paper is organized as follows: Section 2 reviews recent studies related to online exam proctoring systems and the technologies used for detecting cheating in online assessments. Section 3 provides a detailed explanation of the proposed methodology. Section 4 presents the results obtained from the proposed model and compares them with existing techniques. Finally, Section 5 concludes the paper with a summary of findings and insights.

II. LITERATURE REVIEW

The rise of online learning has significantly increased in recent years, prompting researchers to investigate various methods for effectively proctoring online exams. These efforts focus on balancing convenience and efficiency while upholding academic integrity.

In 2022, C. Wang et. al. [6] introduced FAIRSEQ S2T, an extension of FAIRSEQ for speech-to-text tasks like speech recognition and translation. It offers scalable and flexible workflows, covering data pre-processing, model training, and both offline and online inference. The extension includes advanced models based on RNNs, Transformers, and Conformers, with open-source training recipes. Sangeeta and N. Sharma in 2024 [5], outlines existing pedagogical research comparing proctored and unproctored testing, while exploring the technological solutions available for conducting proctored online exams. This study investigates both the human and technical aspects of AIPS, addressing critical issues, and highlights recent innovations that could significantly impact online education and OPS in the future. Banzon et al. [7] examined the use of facial expressions for identification in classrooms, highlighting ethical concerns and proposing guidelines for affect detection. Their essay aims to create a typology of reflexive ethical implications by using a Reflexive Principlism method to track students' emotions through changes in facial expressions. The authors distinguish between the use of this technology in applied education and research, arguing that research applications should be limited until a deeper understanding of the ethical ramifications of affective computing in educational contexts is achieved.

T. Susnjak and T.R. McIntosh in 2024 [8] highlights the threats posed by Large Language Models (LLMs) like ChatGPT to online exam integrity. It introduces a self-reflective strategy that enhances critical thinking and multi-hop reasoning in LLMs when answering complex multimodal exam questions. The strategy was tested on real exam questions and guided ChatGPT to correct answers, showing its strong capability in answering multimodal questions across subjects. The findings challenge previous limitations of LLMs in multimodal reasoning and emphasize the need for advanced proctoring and sophisticated exam designs to prevent AI-driven academic misconduct. J. Deng et. al. in 2019 [9] presents RetinaFace, a highly efficient single-stage face detection model designed for pixel-level localization of faces at different scales. It utilizes a blend of extra-supervised and self-supervised multi-task learning to improve detection accuracy. Additionally, a self-supervised mesh decoder branch is integrated to estimate 3D facial shape information at the pixel level, complementing the supervised components. With its lightweight backbone networks, RetinaFace is capable of real-time operation on a single CPU core for VGA-resolution images.

S. Park et. al. in 2018 [10] proposes an innovative learning-based method for localizing landmarks in the eye region, enabling traditional techniques to compete with modern appearance-based approaches. Despite being trained exclusively on synthetic datasets, the method achieves superior performance in iris localization and eye shape registration on real-world images. The identified landmarks are utilized for iterative model fitting and efficient learning-based gaze estimation. This method demonstrates improved accuracy over existing model-fitting and appearance-based techniques in both person-independent and personalized gaze estimation contexts. N.Ruiz et. al. in 2018 [11] introduced a robust and efficient method for pose estimation, utilizing a multi-loss convolutional neural network trained on 300W-LP, an extensively augmented synthetic dataset. The network predicts intrinsic Euler angles—yaw, pitch, and roll—directly from image intensities by combining binned pose classification with regression techniques. Comprehensive evaluations on prominent in-the-wild pose benchmark datasets highlight its state-of-the-art performance.

Singh, T., et al. in 2024 [12] introduce a robust and efficient framework designed to conduct online assessments while maintaining academic integrity and ensuring fair evaluation of students' knowledge and abilities. To address the challenges of remote exam security, they proposed an advanced Online Exam Proctoring Model utilizing the You Only Look Once (YOLO) algorithm. This automated approach reduces dependency on human proctors, thereby minimizing potential biases and improving scalability for large-scale examinations. By incorporating AI and YOLO, the system facilitates real-time monitoring and immediate detection of suspicious activities, promoting a transparent and reliable assessment process. Ferdosi et al. [13] in 2023 proposed a method to model and categorize student behavior during online exams. Their study outlined a proctoring system tailored for pen-and-paper exams conducted online. This approach tracked head, eye, and lip movements frame by frame, identifying patterns and computing chunk scores (100 frames per chunk) to generate an overall cheating score. The system used MediaPipe to detect facial landmarks and applied K-NN, a high-performing machine learning algorithm, for orientation classification.

Alsabhan [14] introduced a technique in 2023 for identifying cheating among students in higher education using LSTM and other machine learning methods. With a 90% accuracy rate, this approach outperformed three previously referenced models. The framework utilized Long Short-Term Memory (LSTM) networks along with dense layers, a dropout layer, and the Adam optimizer. The high accuracy was credited to the implementation of optimized hyper parameters and a more sophisticated architecture. In 2023, T. Potluri [15] introduced the "Attentive Framework," an AI-based invigilation system for virtual exams. It employs live video analysis with features like appearance verification, multi-person detection, face spoofing prevention, and head pose estimation to detect cheating. Advanced systems also use biometric authentication, keystroke tracking, and 360-degree cameras for enhanced monitoring. In 2022, M. Masud, et al. [16] transformed video data into multivariate time-series by extracting temporal details from each frame, treating dishonesty detection as a sequential time-series problem. In 2021, Nigam, A. et al. [17] conducted a comprehensive survey highlighting that integrating biometric authentication methods into online exams significantly improves their security and reliability. Similarly, Y. Atoum et al. [18] employed advanced techniques such as biometric verification, keystroke analysis, and 360-degree cameras to detect cheating. Additionally, multimedia analytics systems enable automated and continuous proctoring, ensuring the integrity of online assessments

III. METHODOLOGY

We collected a dataset consisting of 18 subjects, with each video ranging from 8 to 20 minutes. We manually labelled different segments of the dataset into 5 categories. The videos were recorded in 720p at an average frame rate of 30 FPS. From this dataset, we selected over 300 chunks of 5-seconds each that met all labelling criteria. We processed webcam input in chunks, which were simultaneously passed through the YOLO-v11 and Silero VAD modules. These chunks were also sent to the tab-switching module for additional processing.

Using the Mediapipe library, facial landmarks were detected, while YOLO-v11 was employed to crop face regions into 128x128 face snippets with 20 pixels of padding. These cropped images were then fed into L2CS-Net for gaze tracking, SixDRepNet for head pose angle estimation, and Mediapipe for 468 3D facial landmark detection. The outputs from these modules, along with the cropped face images, were stored in NumPy arrays and saved in a text file for efficient storage and retrieval.

SYMBOL EXPLANATION

- G: Sequence data containing frames, landmarks, gazes, head poses, and audio.
- M: Pre-trained inference model.
- F: Set of video frames from the sequence.
- L: Set of landmarks extracted from the frames.
- Z: Set of gaze data extracted from the frames.
- H: Set of head poses extracted from the frames.
- N: Set of frame keys where multiple faces are detected.
- ℓ: Predicted label $\in \{0,1,2,3,4\}$.
- $\mathbf{X}$: Concatenated inputs for model inference.
- $\hat{p}$: Model prediction probability.

Algorithm 1 Process Sequence of Frames Algorithm

Input: G: Sequence data (frames, landmarks, gazes, head poses, audio)

M: Pre-trained model for inference

Output: Predicted label $\ell \in \{0,1,2,3,4\}$

Initialize: $F \leftarrow \emptyset, L \leftarrow \emptyset, Z \leftarrow \emptyset, H \leftarrow \emptyset, N \leftarrow \emptyset$

Check Tab Switching:

G.attrs[”tab switch”] = True $\boldsymbol{\ell} = \mathbf{4}$ Tab switching detected

Analyze Audio:

"audio" $\in G$ $a, f_s \leftarrow$ decode audio(G[“audio”]) is audio high energy(a, f_s) $\boldsymbol{\ell} = \mathbf{2}$

Process Video Frames: $k \in G$.keys()

where k starts with “frame ” :

Extract $f_k \leftarrow$ resize($G[k]$[“frame”], (192, 192))

Append $f_k \rightarrow F$

Extract $l_k, z_k, h_k \leftarrow G[k]$[“landmarks, gaze, head pose”]

Append $l_k \rightarrow L, z_k \rightarrow Z, h_k \rightarrow H$

Append $n_k \rightarrow N$ if $G[k]$[“num faces”] > 1

$|N| \geq 44$ $\boldsymbol{\ell} = \mathbf{1}$ Multiple faces detected

Prepare Model Inputs: $F, L, Z, H \leftarrow$ truncate to $(150, F, L, Z, H)$ $\mathbf{X} \leftarrow \{F, L, Z, H\}$

Running the Hybrid CNN-BiLSTM Algorithm:

$\hat{p} \leftarrow (\mathbf{X})$ $\hat{p} < 0.5$ $\boldsymbol{\ell} = \mathbf{0}$ Normal behavior $\boldsymbol{\ell} = \mathbf{3}$ Suspicious

The face images and accompanying data were subsequently input into a Convolutional Neural Network (CNN) for large-scale feature extraction and a Bidirectional Long Short-Term Memory (Bi-LSTM) network to capture temporal patterns in sequential data.

A. *Face Detection and Multi-person detection YOLOv11:* YOLO (You Only Look Once version 11) is an advanced object detection system designed for accurate and efficient face detection, offering real-time performance. Its one-stage architecture allows it to identify and locate faces in images or videos with a single network pass. The framework incorporates an optimized backbone, improved anchor box mechanisms, and sophisticated loss functions, making it effective in detecting small or partially hidden faces. Features such as spatial attention modules and enhanced grid cell predictions enable it to perform well in complex scenarios, including low lighting and occlusion. YOLOv11 is well-suited for real-time applications such as security surveillance, video communication, and augmented reality in diverse fields.

B. *Facial Landmarks detection:* Media Pipe Face Mesh is a powerful tool for real-time face landmark detection, capable of accurately mapping 468 3D facial points. It detects important facial features such as the eyes, nose, lips, and jawline, even in challenging environments with poor lighting or partial obstructions. The framework is lightweight and optimized for performance, making it suitable for various devices, including mobile and web platforms. Its efficiency and accuracy are ideal for applications in facial recognition, augmented reality, emotion detection, and virtual avatar creation. By providing detailed 3D facial coordinates, Media Pipe Face Mesh is an essential resource for industries like gaming, telemedicine, and virtual communication.

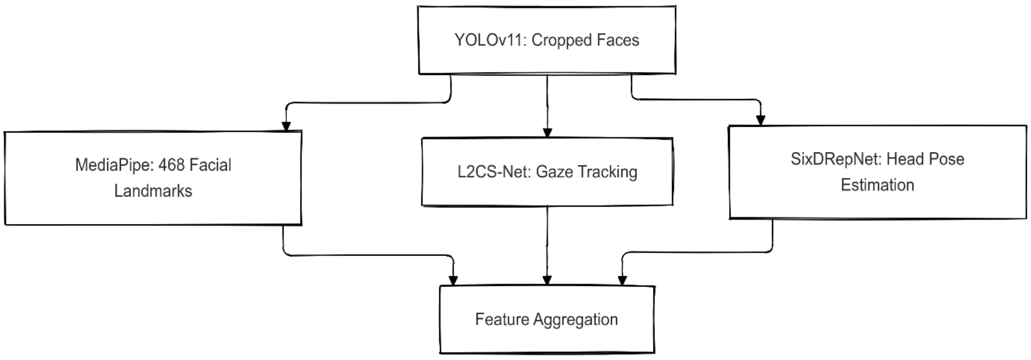

Fig. 2 Modules implemented for data processing

C. *Eye Gaze Tracking:* L2CS-Net (Learned 2D Convolutional Spatial Network) is a sophisticated neural network model designed to provide precise eye gaze tracking. It accurately estimates gaze direction by analyzing the eye's movement and position, even in challenging conditions such as variable lighting and different head orientations. The model leverages deep learning techniques to process eye images and predict gaze based on both spatial and visual characteristics. L2CS-Net is particularly effective at detecting subtle eye movements, making it ideal for real-time gaze tracking applications in fields like human-computer interaction, virtual reality, and driver monitoring systems. Its ability to track gaze with high accuracy and low latency contributes to more engaging and intuitive user experiences, enhancing interfaces and safety features. Given its strong performance, L2CS-Net holds great potential in areas such as cognitive research, interactive gaming, and assistive technologies, where accurate gaze detection is essential for understanding and responding to user intent.

D. *Head Pose Estimation:* SixDRepNet is a sophisticated deep learning model for accurate head pose estimation, capable of predicting 3D head orientation from a single 2D image. By identifying key facial landmarks, it estimates the pitch, yaw, and roll angles with high precision. The model uses a specialized six degrees of freedom

(6DoF) representation, enabling real-time predictions of the head’s full 3D rotation. This technique is particularly useful in fields such as augmented reality, human-computer interaction, and driver monitoring, where head orientation plays a critical role in interaction and safety. SixDRepNet's robustness in handling variations like lighting changes, facial expressions, and partial occlusions makes it ideal for real-world applications. Its accuracy and efficiency make it suitable for both academic research and practical industry use, particularly for real-time head pose tracking.

E. *Voice Activity Detection:* Audio detection involves identifying specific sounds or events in an audio signal through signal processing and machine learning techniques. It analyzes features such as frequency, pitch, and volume to detect sounds, patterns, or activities. Common applications include speech recognition, environmental monitoring, audio surveillance, and music analysis. Advanced systems use deep learning models like CNNs or RNNs to improve detection accuracy, even in noisy conditions. Audio detection is crucial in industries like security (detecting alarms), healthcare (monitoring patient sounds), and entertainment (analyzing music and video). Its real-time capability makes it ideal for interactive systems and smart devices.

IV. RESULTS AND DISCUSSION

The final model weights were saved as a text file, enabling convenient reuse for future applications. Performance metrics, including accuracy, precision, recall, and F1 score, were calculated to assess the model's effectiveness in real-time scenarios. Tab-switching detection is achieved through a front-end JavaScript component utilizing the Visibility Change API, which effectively identifies tab-switching events. Since this component functions independently of the primary model and demonstrates minimal failure rates in modern browsers, we assume it operates with 100% accuracy for evaluation purposes. Although tab-switching events are not explicitly represented in the dataset, their predicted occurrences are factored into the overall accuracy calculation to align with real-world implementation scenarios.

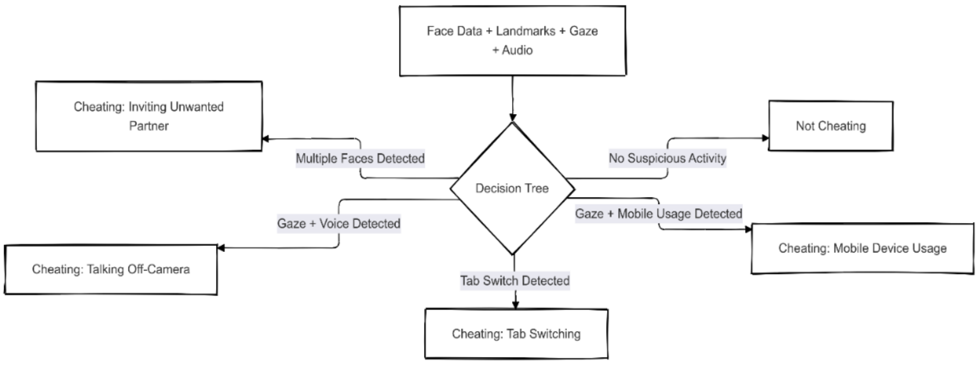

Fig. 3 Decision framework for the proposed methodology

The results in Table 1 demonstrate that the system achieves a high level of reliability and robustness, as indicated by its strong performance metrics, including an accuracy of 0.875, an F-Score of 0.8762, and robust MCC and Kappa values. These findings suggest that the system effectively balances precision and recall while maintaining consistent classification performance, supporting its applicability and effectiveness in the intended domain.

Table 1. Evaluation Metrics for proposed system

Metric	Value
Accuracy	0.875
Error Rate	0.125
F-Score	0.8762
MCC	0.7579
Kappa	0.7511
Sensitivity	0.8214
Specificity	0.9375

AUC	0.8795

Total 18 participants took part in the data collection process, all of whom were actors mimicking the experience of taking an exam. They were asked to exhibit cheating behaviors during the sessions but were not given specific instructions on the types or methods of cheating to perform. One drawback of this approach is the potential presence of unnatural behaviors resulting from the acting. The sample outputs displayed in this study consist of frames taken from video sequences in the generated dataset. The results of the face detection and landmark localization algorithm are illustrated in Fig. 4.

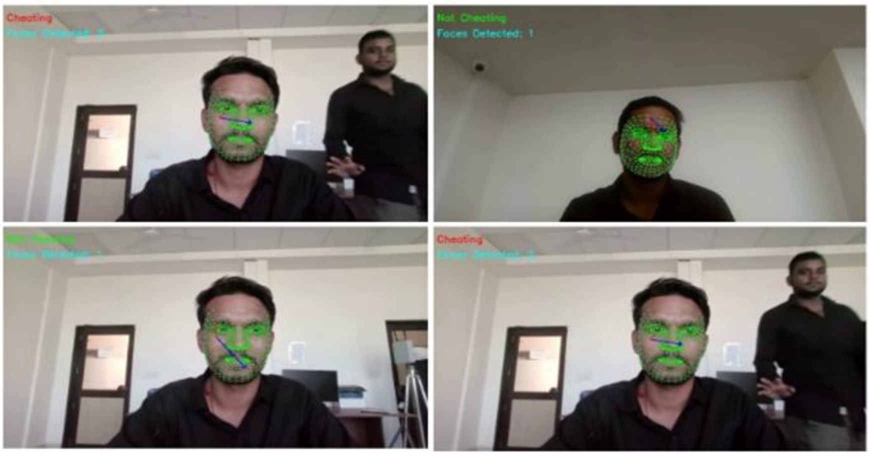

Fig.4 Sample Dataset Output Screenshots

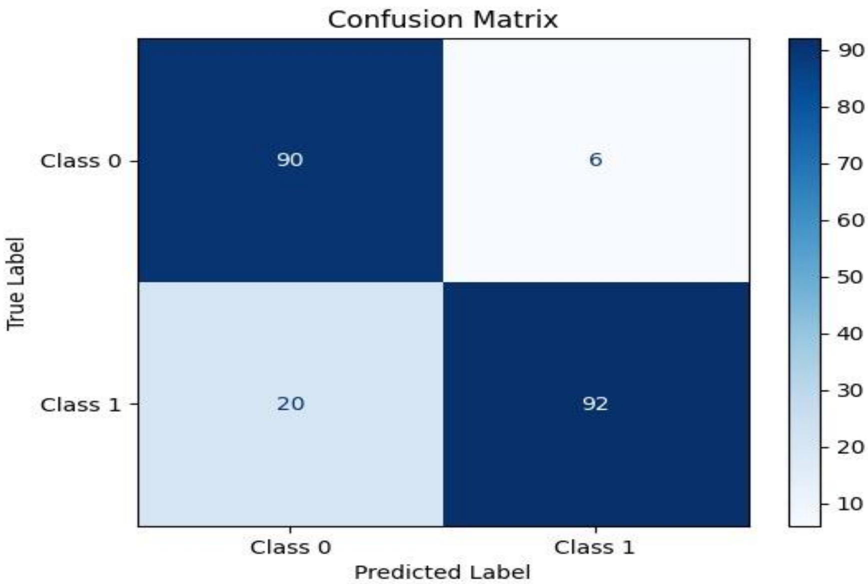

Fig. 5 Confusion matrix for Predicted Label and True Label

The confusion matrix presented in Figure 5 evaluates the model's performance in detecting cheating and non-cheating behaviors. It shows that the model correctly identified 90 instances of non-cheating (true negatives) and 92 instances of cheating (true positives). However, there were 6 instances where non-cheating behavior was misclassified as cheating (false positives) and 20 instances where cheating behavior was incorrectly classified as non-cheating (false negatives). This analysis, based on a total of 208 samples, indicates that while the model performs well in distinguishing between the two classes, there is potential for improvement in reducing false positives and false negatives.

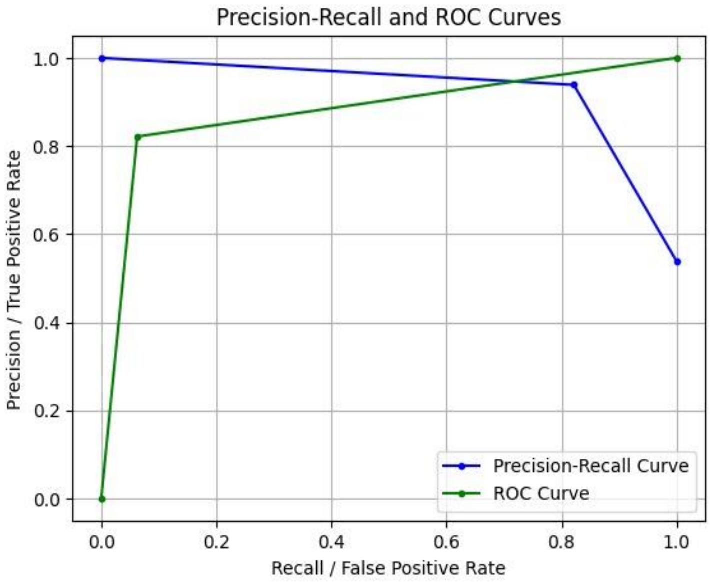

Fig. 6 Comparison of Precision Recall Curve and ROC Curve

The figure 6 above displays the Precision-Recall curve and the Receiver Operating Characteristic (ROC) curve, providing insights into the model's performance in terms of classification effectiveness. The Precision-Recall curve (blue) exhibits an initial sharp increase, followed by stabilization, indicating high precision and recall in detecting the positive class. This suggests that, for most thresholds, the model effectively distinguishes positive instances, maintaining a strong balance between precision and recall. In contrast, the ROC curve (green) shows a high True Positive Rate (TPR) with an associated False Positive Rate (FPR), reflecting the model's ability to discriminate between the positive and negative classes. Both curves suggest that the model performs well across various thresholds, with the Precision-Recall curve being particularly informative in scenarios where class imbalance is a concern, as it focuses on the performance regarding the positive class. These visualizations confirm that the model is adept at identifying the positive class while minimizing false positives and negatives.

V. CONCLUSION

This research presents an effective model for detecting behavioral anomalies using a multimodal dataset. The experimental results demonstrate a high degree of performance, with an overall accuracy of 87.5% and an F-Score of 87.62%, indicating a strong balance between precision and recall. Furthermore, the Matthews Correlation Coefficient (MCC) of 0.7579 and Cohen's Kappa of 0.7511 underscore the reliability of the model across diverse data distributions. The model exhibits excellent specificity (93.75%), affirming its ability to correctly identify non-anomalous behaviors, while achieving a sensitivity of 82.14%, reflecting its capacity to detect anomalous cases accurately. The Area Under the Curve (AUC) score of 0.8795 further corroborates the robustness of the classifier in distinguishing between the behavioral classes. These results validate the efficacy of incorporating multimodal features such as audio, gaze, and head pose, coupled with advanced neural architectures. The findings not only establish the utility of this model in practical proctoring systems but also highlight areas for future improvement, such as addressing edge cases where sensitivity could be enhanced. This work provides a foundation for further exploration of multimodal behavior analysis in real-world applications.

STATEMENTS AND DECLARATIONS

CONFLICT OF INTEREST: The authors confirm that they have no financial, personal, or professional conflicts of interest related to this study.

FUNDING: This study did not receive financial support from any funding organization.

DATASET AVALABILITY STATEMENT: The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the anonymous volunteers who participated in the collection of our REP database. Additionally, they extend their thanks to B.H.V.S.P. Subrahmanyam, K. Kant, and S. Saraswat for their valuable contributions to the study of human behavior analysis during e-exam proctoring.

REFERENCES

[1] URL: https://nces.ed.gov/programs/digest/d23/tables/dt23_311.15.asp

[2] URL: Online Exam Proctoring Market Growth Report 2021 to 2031

[3] A. Nigam, R. Pasricha, T. Singh and P. Churi, "A Systematic Review on AI-based Proctoring Systems: Past, Present and Future" Education and Information Technologies, vol. 26, june 2021, pp. 6421–6445, https://doi.org/10.1007/s10639-021-10597-x

[4] Y. Atoum, L.Chen, A.X. Liu, S.D.H. Hsu and X. Liu, "Automated Online Exam Proctoring" IEEE Transaction on Multimedia, Dec 2015, pp. 1-15.

[5] Sangeeta and N. Sharma, "Role of AI in Online Examination Proctoring Systems" book chapter in Applications of Artificial Intelligence in the Interne of Things: Today's and Tomorrow's World in Cambridge Scholars Publishing, 2024, pp. 301-315.

[6] C. Wang, Y. Tang , X. Ma, A. Wu, S. Popuri, D. Okhonko and J. Pino, "FAIRSEQ S2T: Fast Speech-to-Text Modeling with FAIRSEQ" arXiv preprint arXiv:2010.05171, june 2022.

[7] A.M. Banzon, J. Beever,and M. Taub, " Facial Expression Recognition in Classrooms: Ethical Considerations and Proposed Guidelines for Affect Detection in Educational Settings" IEEE Transactions on Affective Computing, 2023, pp. 93-104.

[8] T. Susnjak and T.R. McIntosh, " ChatGPT: The end of online exam integrity?" *Education Sciences*, *14*(6), p.656, 2024, https://doi.org/10.3390/educsci14060656.

[9] J. Deng, J. Guo,Y. Zhou, J. Yu, I. Kotsia and S. Zafeiriou, "RetinaFace: Single-stage Dense Face Localisation in the Wild" arXiv preprint arXiv:1905.00641, may 2019.

[10] S. Park, X. Zhang, A. Bulling and O. Hilliges, "Learning to Find Eye Region Landmarks for Remote Gaze Estimation in Unconstrained Settings" In Proceedings of the 2018 ACM symposium on eye tracking research & applications, pp. 1-10, june 2018.

[11] N. Ruiz, E. Chong and J. M. Rehg, "Fine-Grained Head Pose Estimation Without Keypoints" In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 2074-2083, 2018.

[12] T. Singh, R.R. Nair, T. Babu and P. Duraisamy, "Enhancing Academic Integrity in Online Assessments: Introducing an Effective Online Exam Proctoring Model using YOLO" Procedia Computer Science, 235, pp.-1399-1408, 2024.

[13] B.J. Ferdosi, M. Rahman, A.M. Sakib and T. Helaly, "Modeling and Classification of the Behavioral Patterns of Students Participating in Online Examination" Human Behavior and Emerging Technologies, 2023.

[14] W. Alsabhan, "Student cheating detection in higher education by implementing machine learning and LSTM techniques" Sensors, 23(8), p.4149, 2023

[15] T. Potluri, "An automated online proctoring system using attentive-net to assess student mischievous behavior" Multimedia Tools and Applications, pp. 1-30, 2023

[16] M.M. Masud, K. Hayawi, S.S. Mathew, T. Michael, and M. E. Barachi, "Smart online exam proctoring assist for cheating detection" In Advanced Data Mining and Applications, Chem: Springer International Publishing, pp. 118-132, Jan 2022

[17] A. Nigam, R. Pasricha, T. Singh, and P. Churi, "A systematic review on AI-based proctoring systems: Past, present and future" Education and Information Technologies, 26(5), 6421–6445, 2021. https://doi.org/10.1007/s10639-021-10597-x

[18] Y. Atoum, L. Chen, A.X. Liu, S.D.H. Hsu, and X. Liu, "Automated online exam proctoring" IEEE Transactions on Multimedia, 19(7), pp. 1609–1624, 2017. https://doi.org/10.1109/TMM.2017.2656064