VIRTUAL VIGILANCE USING CNN

Urvi Padelkar
Computer Science and Engineering Data Science
PCT's A. P. Shah Institute of Technology
Thane, 400615, India
Email: urvipadelkar449@apsit.edu.in

Pranjal Desai
Computer Science and Engineering Data Science
PCT's A. P. Shah Institute of Technology
Thane, 400615, India
Email: pranjaldesai457 @apsit.edu.in

Dhanashree Kasar
Computer Science and Engineering Data Science
PCT's A. P. Shah Institute of Technology
Thane ,400615, India
Email: dhanashreekasar435@apsit.edu.in

Vineet Mhatre
Computer Science and Engineering Data Science
PCT's A. P. Shah Institute of Technology Thane, 400615, India
Email: vineetmhatre @apsit.edu.in

Poonam Pangharkar
PCT's A. P. Shah Institute of Technology
Thane, 400615, India Email: @apsit.edu.in

Abstract - The widespread adoption of remote learning has led to a surge in online assessments, raising concerns about maintaining academic integrity. Traditional methods of proctoring are not feasible in virtual settings, necessitating innovative solutions. Virtual Vigilance have emerged as a promising approach, leveraging computer vision and machine learning to monitor examinee behaviour. This paper presents a comprehensive analysis of AI proctoring systems, focusing on facial recognition, eye tracking, lip movement detection, and head movement analysis. The proposed system integrates advanced algorithms from OpenCV and DeepFace libraries to ensure the credibility of online examinations. Through real-time monitoring and alerting, the system detects and prevents cheating behaviours, thereby upholding the integrity of online tests.

Keywords – Remote learning, Integrity, Facial recognition, eye tracking, lip movement detection, head movement analysis, OpenCV, DeepFace, real-time monitoring, cheating prevention.

I. Introduction

The rise of remote learning has brought about unprecedented convenience for students, allowing them to access educational content from anywhere in the world. However, this shift has also raised concerns regarding academic integrity during online assessments. These systems have garnered attention as a vital tool in addressing the challenges of remote learning and online assessment. With the proliferation of virtual classrooms and distance education, the need to uphold academic integrity has become essential. Traditional methods of proctoring are impractical in virtual settings, prompting the need for innovative solutions. These systems typically incorporate a range of monitoring techniques, such as facial recognition, eye tracking, lip movement detection, and head movement analysis, to accurately assess examinee behaviour during exams.

II. LITERATURE REVIEW

There have been giant leaps in the field of education in the past 1–2 years. Schools and colleges are transitioning online to provide more resources to their students. The COVID-19 pandemic has provided students more opportunities to learn and improve themselves at their own pace. Online proctoring services (part of assessment) are also on the rise, and AI-based proctoring systems (henceforth called as AIPS) have taken the market by storm. Face recognition is an important research problem spanning numerous fields and disciplines. The author proposed collecting facial profiles as curves, finding their norm, and then classifying other profiles by their deviations from the norm. This classification is multi-modal, i.e. resulting in a vector of independent measures that could be compared with other vectors in a database [1].

Eye tracking is the process of tracking the movement of the eyes to know exactly where and for how long a person is looking. The primary purpose of eye movement is to direct the eyes towards the targeted object and keep it at the centre of the fovea to provide a clear vision of the object. Eye tracking is used in various research fields such as cognitive science, psychology, neurology, engineering, medicine and marketing, to mention some. Human-computer interaction is another example of application it is beneficial for disabled people to interact with a computer through the gaze. Eye tracking can also be used to monitor and control automobile drivers. It is thus highly interdisciplinary and used in various fields, which is also reflected in how eye-tracking hardware and software have been developed over the years. To extract useful information, the raw eye movements are typically converted into so-called events. This process is named event detection [2].

Recognition of Lip movements has become one of the most challenging tasks and has crucial applications in the contemporary scenario. It is the recognition of the speech uttered by individual using visual cues. Visual interpretation of lip movement is especially useful in scenarios like video surveillance, where auditory signals are either not available or too noisy for interpretation. It is also useful for hearing-impaired individuals where the audio signal is of no use. Many developments have taken place in this nascent field using various deep learning-based techniques[3].

The increased popularity of the wide range of applications of which head movement detection is a part, such as assistive technology, teleconferencing and virtual reality, have increased the size of research aiming to provide robust and effective techniques of real-time head movement detection and tracking. One approach for head movement detection is computer vision-based. Liu et al. introduced a video-based technique for estimating the head pose and used it in a good image processing application for a real-world problem; attention recognition for drivers. It estimates the relative pose between adjacent views in subsequent video frames. Scale-Invariant Feature Transform (SIFT) descriptors are used in matching the corresponding feature points between two adjacent views. After matching the corresponding feature points, the relative pose angle is found using two-view geometry. With this mathematical solution, which can be applied in the image-processing field in general, the x, y, and z coordinates of the head position are determined. The accuracy and performance of the algorithm were not highlighted in the work and thus more work is needed to prove this algorithm to be applicable in real applications[4].

Human face detection has been a challenging issue in the areas of image processing and patter recognition. A new human face detection algorithm by primitive Haar cascade algorithm combined with three additional weak classifiers is proposed in this paper. The three weak classifiers are based on skin hue histogram matching, eyes detection and mouth detection. First, images of people are processed by a primitive Haar cascade classifier, nearly without wrong human face rejection (very low rate of false negative) but with some wrong acceptance (false positive). Secondly, to get rid of these wrongly accepted non-human faces, a weak classifier based on face-skin hue histogram matching is applied and a majority of non-human faces are removed. Next, another weak classifier based on eye detection is appended and some residual non-human faces are determined and rejected. Finally, a mouth detection operation is utilized on the remaining non-human faces and the false positive rate is further decreased. With the help of OpenCV, test results on images of people under different occlusions and illuminations and some degree of orientations and rotations, in both training set and test set show that the proposed algorithm is effective and achieves state-of-the-art performance. Furthermore, it is efficient because of its easiness and simplicity of implementation [5].

The Viola-Jones method, a cascade classifier, pioneered by Viola and Jones, is highly effective for face detection and applicable to various object detection tasks with rigid structures. It utilizes Haar-like features, which are efficiently computed from integral images. While Local Binary Patterns (LBP) is another viable feature, this paper focuses solely on Haar-like features. The Viola-Jones detector employs AdaBoost, forming a rejection cascade comprising multi-tree AdaBoosted classifiers. Despite occasional false positives, with a sufficient number of nodes, the method achieves an impressively high face detection rate while maintaining a low false positive rate.[6]

III. PROPOSED SYSTEM

The proposed AI proctoring system is designed to monitor and analyse examinee behaviour during remote examinations using advanced computer vision techniques and machine learning algorithms. It comprises several modules that work collaboratively to detect and flag potential instances of academic misconduct, ensuring the integrity and fairness of online tests.

- Algorithm Overview:

The system utilizes deep learning models provided by the DeepFace library for facial recognition and verification of examinees' identities. Facial features extracted from video frames captured by the webcam are compared against reference images or templates associated with authorized examinees. Eye tracking algorithms, implemented using OpenCV, monitor the movement and gaze patterns of examinees' eyes in real-time to detect suspicious behaviours related to visual attention. Lip movement analysis techniques, including Gaussian blur and contour detection, are applied to examine lip movements and expressions for signs of verbal communication or reading aloud exam content. Head movement detection algorithms analyze the motion and posture of examinees' heads throughout the examination to identify signs of restlessness or physical manipulation.

IV. RESULT AND ANALYSIS

The AI-powered monitoring system represents a sophisticated integration of computer vision algorithms and facial recognition capabilities, aimed at real-time analysis of users' facial expressions and movements. Here's an in-depth exploration of the theoretical underpinnings and the associated calculations:

Haarcascade Classifier:

The Haar Cascade classifier is a machine learning-based approach used for object detection. It works by analysing features.

1.Feature Selection: Haar features are rectangular patterns used to represent localized image features. These features capture contrast differences in adjacent regions.

2.Training: The classifier is trained using positive and negative images. Positive images contain the object of interest, while negative images don't. The algorithm learns to distinguish between these images by adjusting weights on different features.

3. Detection: During detection, the classifier slides over the image at different scales and positions, applying each stage of the cascade sequentially. If a region passes all stages, it's classified as containing the object of interest.

1. Facial Recognition and Analysis:

DeepFace Integration: DeepFace, a deep learning-based facial recognition library, is employed to compare the captured facial image with a reference image. The verification process involves calculating a similarity score between the two images using deep neural networks. Threading is used to run the facial recognition task in parallel with frame capturing and display, ensuring smooth real-time processing. DeepFace internally calculates the cosine similarity or Euclidean distance between the feature embeddings of the two images to determine the similarity score.

Fig.1 Face Recognition

2. Eye Movement Monitoring:

Threaded Eye Tracking: Eye movement tracking involves detecting and monitoring the movement of eyes within the captured frame. This process is executed using Haar cascade classifiers for eye detection and optical flow algorithms for tracking eye movements.

Formula (Optical Flow): The mean displacement D of key points between two consecutive frames can be calculated as:

$D=1/N\sum_{i=1}^{N}\sqrt{(x_{inew}-x_{iold})^2+(y_{inew}-y_{iold})^2}$

Fig.2 Eye Movement Detection

3. Lip Movement Detection:

Lip movement detection involves analysing changes in lip contour and shape within the captured frame. Gaussian blur is applied to reduce noise, followed by contour detection to identify lip movements.

Formula (Contour Area): The key function used in the code is cv2.countourArea(contour) which calculates the contour area of detected regions in the image.

A=cv2.contourArea(contour)

A represents the contour area, cv2.countourArea() is the function provided by OpenCV to calculate the contour area of detected regions in the image and contour refers to the detected contour of the lips in the image.

Calculation: After applying Gaussian blur and contour detection, the contour areas of detected lips are computed. If the contour area exceeds a predefined threshold, it indicates significant lip movement.

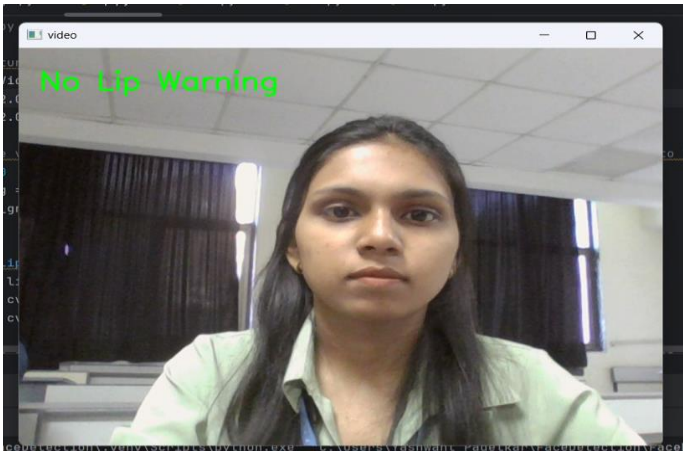

Fig.3 Lip Movement Detection

4. Head Movement Analysis: Head movement analysis utilizes optical flow estimation to calculate the displacement of key points (e.g., facial landmarks) between consecutive frames. The magnitude of optical flow vectors indicates the extent of head movement.

Formula (Motion Percentage): The motion percentage P can be calculated as the ratio of the number of key points with motion to the total number of key points:

P = Number of key points with motion/ Total number of key points ×100%

Calculation: By computing the motion percentage, the system assesses the overall movement of key points, providing insights into head movement.

Fig.4.Head Movement Detection

5. Real-time Monitoring and Alerting: The system continuously captures frames from the webcam and processes them in real-time to monitor various facial expressions and movements. Based on the analysis results, the system generates alerts displayed directly on the video feed, providing immediate feedback to users. The thresholds for triggering alerts are predefined based on empirical data and domain knowledge. By incorporating these theoretical concepts and associated calculations, the AI-powered monitoring system achieves accurate and responsive analysis of facial expressions and movements, thereby enhancing user safety and performance in diverse applications.

V. CONCLUSION

In conclusion, Virtual Vigilance is AI-powered proctoring systems represent a significant advancement in ensuring the integrity of online assessments. By leveraging computer vision and machine learning techniques, these systems can effectively monitor examinee behaviour in real-time. The integration of facial recognition, eye tracking, lip movement detection, and head movement analysis enables comprehensive monitoring of examinee activities during online examinations. The proposed system demonstrates promising results in detecting and preventing cheating behaviours, thereby maintaining the credibility of online tests. However, further research and development are needed to enhance the accuracy and robustness of AI proctoring systems and address challenges such as privacy concerns and algorithm bias.

VI. REFERENCES

[1] Nigam, A., Pasricha, R., Singh, T. et al. A Systematic Review on AI-based Proctoring Systems: Past, Present and Future. Educ Inf Technol 26, 6421–6445 (2021). https://doi.org/10.1007/s10639-021-10597-x

[2] Birawo, B.; Kasprowski, P. Review and Evaluation of Eye Movement Event Detection Algorithms. Sensors 2022, 22, 8810

[3] Srilakshmi, Kanagala, and R. Karthik. "A Novel Method for Lip Movement Detection using Deep Neural Network." Journal of Scientific & Industrial Research 81.06 (2022): 643-650.

[4] A. Al-Rahayfeh and M. Faezipour, "Eye Tracking and Head Movement Detection: A State-of-Art Survey," in IEEE Journal of Translational Engineering in Health and Medicine, vol. 1, pp. 2100212 -2100212, 2013, Art no. 2100212, doi: 10.1109/JTEHM.2013.2289879.

[5] L. Cuimei, Q. Zhiliang, J. Nan and W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), Yangzhou, China, 2017, pp. 483-487, doi: 10.1109/ICEMI.2017.8265863. s

[6] Li Cuimei1 , Qi Zhiliang 2 , Jia Nan 2 , Wu Jianhua 2* 1.School of Communication and Electronics, Jiangxi Science & Technology Normal University Nanchang 330013, china