How to Cite:
Gopalakrishnan, K., Dhiyaneshwaran, N., & Yugesh, P. (2022). Online proctoring system using image processing and machine learning. *International Journal of Health Sciences*, *6*(S5), 891–899.
https://doi.org/10.53730/ijhs.v6nS5.8777

Online proctoring system using image processing and machine learning

Dr. K. Gopalakrishnan AP/ECE
Department of ECE, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India.
Email: kgk1969@mepcoeng.ac.in

N. Dhiyaneshwaran
Department of ECE, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India.

P. Yugesh
Department of ECE, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India.

***Abstract*---**Different forms of remote education and massive online open courses are gaining reputation. The skill to proctor the online examination is a main essential factor for the scalability for promoting the students for next stage education. Existing manual monitoring is the most approaching method in education either by visually monitoring or by physically accompanying test takers to the examination centre and monitoring them. Learner's identity verification and proctoring of online examinations is one of the main challenges in online learning systems. The migration and implementation of the online exam have been accelerated during the pandemic COVID-19. So the existing systems need a safe mechanism to authenticate and proctor online students. In this paper, we propose a system for providing the solution for authentication and proctor by using the different biometric technologies. The performance of the proposed system is evaluated by using real time videos with different scenario. The experimental results provide an improved accuracy than existing research works.

***Keywords*---**Continuous face recognition, Gaze tracking, Mouth open or close.

Manuscript submitted: 18 March 2022, Manuscript revised: 9 April 2022, Accepted for publication: 27 May 2022

Introduction

Due to COVID-19, Most of the Out-dated educational organizations proposing face to face teaching, and these institutes had to push into the fully online educational model forced by the pandemic. Today's educational institutions reach their expand significantly due the potential provided by Massive Open Online Courses (MOOCs). Recent survey says that more than 7.1 million students undergo online education with a different medium. (https://link.springer.com/article/10.1007/s40031-021-00581-x). The digital approaches motivate the students to learn like Animations (54.7%), Whiteboard and pen (36.5%), PowerPoint Examinations are the utmost perilous component of any educational institution. They should reproduce the real learning in order to hold their worth to the student society. According to the survey (https://www.educationworld.in/73-of-students-cheat-during-online-exams-survey/), 73% of students are cheating in online exams. The recent survey showed that 36% undergraduate and 24% postgraduate students are those who paraphrase from the internet without footnoting, 38% undergraduate and 25% postgraduate students are those who paraphrase from the book without footnoting, and 7% undergraduate students are those who copy word to word without citation. When the exams are conducted in conventional and presentation (37.69%), Digital pen and slate (33.63%). In offline examination centres, the candidates are monitored by manually. In contrast, there is no human proctor for students taking online exams. In this paper, the author compresses various detection methods that are not shown in the latter proposed methods. The various detection techniques are Continuous Face Recognition, Gaze Tracking, Mouth Open or Close. The video information is collected from Camera. By using Microphone the audio information has collected. The data's from screen sharing also collected. These are three foremost data that are collected during online education. Of these, video information is still not used to its full potential. In the current day online examination monitoring systems, the face process alone employed. The features from videos are collected from webcam and it is used for monitoring. But this process provides minimum accuracy. Hence, additional features like estimation of head pose, monitoring close and open condition of mouth, monitoring the movement of eyeball, detecting the presence of other person, presence of mobile phones, and identifying face spoofing are required for improving the efficiency. Existing system monitor only face movements and sound information. In this paper, we give additional features like, Gaze Tracking, and monitoring the Mouth open and close. The research on online proctoring systems related work is discussed in Section II. The suggested system's approach and design are described in Section III. The system performance is evaluated in Section IV. Section V wraps up the report and looks ahead to future projects.

Related Works

Many aspects of our life have been disrupted by the COVID-19 epidemic. To reduce human infection rates, governments all over the world have regulated physical interaction for education, work, and recreation. As a result of this predicament, education and many other physical activities and enterprises have been compelled to relocate from real to virtual environments. Around the world, school, university, and other forms of education and training have shifted to online learning. Many obstacles, however, are preventing widespread adoption by governments and the general public. Due to the lack of physical presence, monitoring attendance and students throughout courses, particularly during tests, is a big difficulty for online systems. It is necessary to create methods and technology that will give reliable instruments for detecting unfair, immoral, and unlawful behaviour in courses and test. The current literature in this area is minimal, with the majority of software coming from commercial firms that offer restricted and "non-open" software solutions. To provide innovation, diversity, and depth to the online learning software systems sector, several open-source tools and activities are required. This article offers a revolutionary online proctoring system that use deep learning to continuously proctor physical locations without the requirement for a physical proctor's presence. Recent years, the requirement of online education is increasing expressively. Researchers have proposed various online proctoring methods to make learning easy and convenient way. The methods have been categorized into three methods. They are No proctoring, online human monitoring, partial-automated machine proctoring.

In No proctoring system the cheating is minimized in various ways. Online exam proctoring system with visual and aural sensors is used for academic integrity and upkeep. Online human proctoring is one of the common approaches to monitor test-takers. But one of the downsides is that it is very costly to hire employees to monitor the test takers. The drawback is that a single camera cannot see the all the subject sees. This will results, faulty decision in any cheating strategies. Recently researchers proposed many methods to overcome the needs for online exam proctoring. Fully automated online exam proctoring system with visual and audio sensors for the purpose of academic integrity and maintenance which was proposed by Yousef Atoum [1]. This paper aims to develop a multimedia analysis system to detect a wide variety of cheating behaviours during an online exam session. Asep Hadian S. G. and Yoanes Bandung proposed a method to develop the robustness for pose and lighting variations by applying an incremental training process. In this system the training data sets are obtained. From the data sets learning online lecture session are monitored [2].

Makes use of transfer learning proposed by Ashwinkumar J S, to realize deep learning and it combines three models namely YOLO (for fraudulent object and multi-person detection), MPGazell (for abnormal gaze detection) and VGG16 (for Face recognition) [3]. A Systematic Review on Artificial Intelligence -Based Proctoring System, the author compares the methodologies used for online proctoring and concludes with the issues that the system is going to face in the future [4].

Methodology and Design

The Proposed Framework

The proposed framework has combined the following features – continuous face detection and recognition, mouth open or close, gaze tracking for the entire duration of the exam[5] - [7]. In the preparation phase (before taking the exam), the test taker has to submit 15 – 20 images with their registered number. In the exam phase (during exam duration), the proposed system converts the live video into frames and analyses the frames for suspicious behaviour [8]. After the exam, the examiner receives a score based on the test taker's behaviour, the video of the entire exam, and a report with details of each frame that is analysed. Based on the above details the examiner can decide the credibility of the test taker.

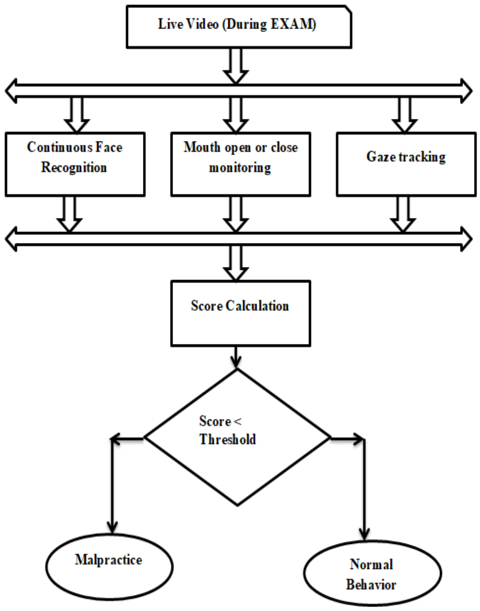


Figure.2 Flow Diagram for the proposed work

Face Detection

To ensure that the test taker remains present during the exam, face detection is essential. The suggested system uses the haar [9] cascade algorithm to recognize faces. A Haar-like feature adds the pixel intensities in consecutive rectangular sections of a detection window at a specified position and calculates the difference between the sums. This differentiation is then utilized to divide picture subsections into categories [9]. Haar cascade algorithm is a machine-learning-based technique. It implicates training a cascade function with the help of huge numbers of positive and negative photos [10].

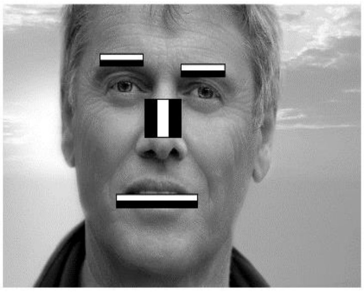


Fig 3 Haar Cascade

Face Recognition

Face recognition is the most prevalent biometric option for internet authentication. Intel established OpenCV in 1999, and it is a well-known computer vision library. The three facial recognition algorithms implemented by OpenCV are Eigenface, Fisherface, and LBPH (Local Binary Patterns Histograms).[11].
For face recognition, the proposed system employs an LBPH classifier. During the training phase, a python script traverses and trains 15-20 photos of the test taker. A YAML file is used to hold the learned model. The LBPH classifier from the OpenCV package is utilized in the recognition step[12].

LBPH Classifier

Applying the LBP operation: The first stage in the LBPH's computational process is to build an intermediate picture that better represents the original image by highlighting the facial features [13]. To do so, the method employs a sliding window idea based on the radius and neighbours parameters.

Divide the image created in the previous step into numerous grids to extract the histograms. Then, for each section, extract the histogram. Then, to make a new, larger histogram, we must concatenate each histogram. The features of the original image are represented by the final histogram [14].
This feature ensures that the person taking the exam is the person registered for the exam.

Mouth open or close

The suggested system analyses the movement of the mouth and determines whether it is open or closed to detect if the test taker is talking or not. The suggested method would do this by utilising a pre-trained network from the dlib library that can detect 68 critical locations on the face[15]. Dlib is employed because, unlike a CNN model, it can make predictions in real time. 68 landmarks - The dlib library's pre-trained facial landmark detector is utilised to estimate the position of 68 (x, y)-coordinates that correspond to facial structures. 300-W, XM2VTS, FRGC Ver2, LFPW, HELEN, AFW, and IBUG datasets were used to train this detector [16].

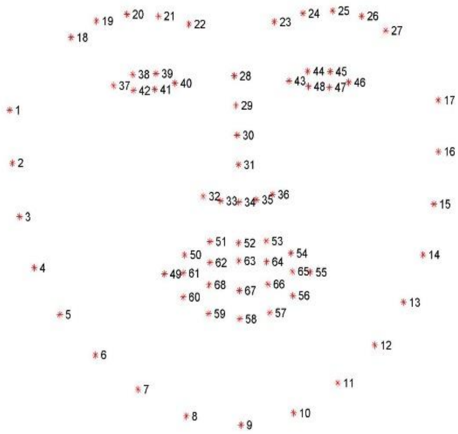


Figure. 3 68 Landmarks

Mouth Aspect Ratio (MAR) is calculated to determine whether the mouth is open or close.

Mouth Aspect Ratio (MAR):

A = Euclidean distance (51, 59)

B = Euclidean distance (53, 57)

C = Euclidean distance (49, 55)

$$MAR\frac{(A+B)}{(2.0*C)} \qquad (1)$$

MAR THRESHOLD = 0.70

When MAR > MAR THRESHOLD:

Mouth is open

When MAR > MAR THRESHOLD:

Mouth is closed

This feature ensures that the test taker is not taking to anyone during the exam.

Gaze Tracking

Gaze tracking feature estimates the coordinates of the pupil and it tracks the movement of eyeball. This is done to ensure that the test taker is seeing only the computer screen during the exam. This feature makes use of Dlib's facial key points once more [17].

The module for this feature gives two values: one is a number between 0.0 and 1.0 that indicates the horizontal direction of the gaze, and the other is a number between 0.0 and 1.0 that shows the vertical direction of the gaze. The extreme right has a value of 0.0, the middle has a value of 0.5, and the extreme left has a value of 1.0. The other is a value between 0.0 and 1.0 that represents the gaze's vertical orientation. 0.0 is at the extreme top, 0.5 is in the middle, and 1.0 is at the extreme bottom. With these values, the proposed system detects whether the test taker is looking left, centre, right or down. This feature ensures that the test taker is not referring to any other study material during the examination [18].

Score Calculation

As an output, the examiner gets the video for the entire exam duration, a report consisting of the details of each frame analysed, and a score that is calculated from the features mentioned. For score calculation, the proposed system uses percent and average measures. After the exam, we get the following – total number of frames, number of frames the face is recognized, number of frames the mouth is open, number of frames the mouth is closed, number of frames the test taker is looking centre, left, right and down. From these numbers, we calculated recognized frames percentage, mouth closed frames percentage, and looking centre frames percentage. To obtain the final score we calculated the average of recognized frames percentage, mouth closed frames percentage, and looking centre frames percentage [19].

Result and Discussion

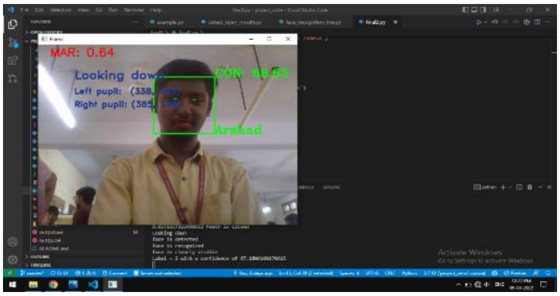
We tested our proposed system with 20 one-minute video clips, of which 10 videos contain malpractice instance, other 10 videos doesn't contain any malpractice instances. After testing we got the following result,

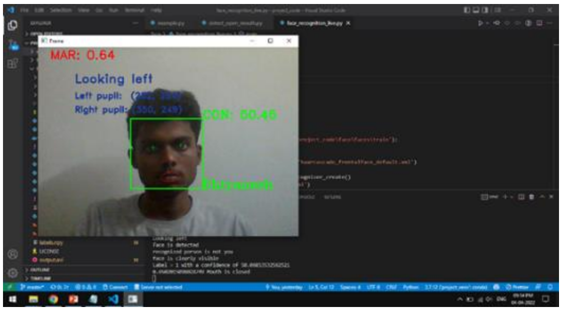
Table 1
Confusion matrix for our proposed system

A \ P	NO	YES	
NO	10	0	10
YES	2	8	10
	12	8	

True positive (TP) – The cases which are predicted as yes, and also the real case is also yes.
True negative (TN) – The cases which are predicted as no, and also the real case is also no.
False positive (FP) – The cases which are predicted as yes, but the real case is no.
False negative (FN) – The cases which are predicted as no, but the real case is yes.
TP = 8, TN = 10, FP = 0, FN = 2

Accuracy = (TP+TN)/Total. Accuracy tells how often the classifier is correct.
Accuracy for the proposed system = (10+8)/20 = 0.9
With the proposed system, the examiner gets the recorded video of the exam. The frames from that video are shown below:





Conclusion and Future Work

The proposed system has accuracy of 90%. To improve accuracy, some more features can be added. Features like multiple person detection and mobile phone detection can be added. The proposed system doesn't have a user interface. For more comfortability user interface can be added in future.

References

[1] Yousef Atoum, L iping Chen, Alex X. Liu, Stephen D. H. Hsu, and Xiaoming Liu," Automated Online Exam Proctoring", by IEEE transaction medium, Jan 2017.

[2] Asep Hadian S. G. and Yoanes Bandung, "A Design of Continuous User Verification for Online Exam Proctoring on M-Learning", by IEEE International Conference on Electrical Engineering and Informatics, in July 2019.

[3] Ashwinkumar J S, Harshavardhini Saravana Kumaran, Sivakarthikeyan U, Konjeti P B V Rajesh, Lavanya R, "Deep Learning Based Approach for Facilitating Online Proctoring using Transfer Learning", by IEEE International Conference on Computer, Communication and Signal Processing, in June 2021.

[4] Aditya Nigam, Rhitvik Pasricha, Tarishi Singh and Prathamesh Churi, "A Systematic Review on AI-based Proctoring Systems: Past, Present and Future" in Springer, in June 2021.

[5] Istiak Ahmad, Fahad AlQurashi, Ehab Abozinadah, Rashid Mehmood, "A Novel Deep Learning-based Online Proctoring System using Face Recognition, Eye Blinking, and Object Detection Techniques", International Journal of Advanced Computer Science and Applications, October 2021.

[6] R. Mehmood, F. Alam, N. N. Albogami, I. Katib, A. Albeshri, and S. M. Altowaijri, “UTiLearn: A Personalised Ubiquitous Teaching and Learning System for Smart Societies,” IEEE Access, 2017.
[7] L. K. Musambo and J. Phiri, “Student facial authentication model based on OpenCV’s object detection method and QR code for Zambian higher institutions of learning,” Int. J. Adv. Comput. Sci. Appl. Jan. 2018.
[8] ProctorU. (2021). The Leading Proctoring Solution for Online Exams.[Online]. Available: https://www.proctoru.com/
[9] Examity. (2021). Better Test Integrity. [Online]. Available: https://examity.com/
[10] PSIOnline. (2021). Certification Testing Services and Programs. [Online].Available: https://www.psionline.com/en-gb/certification/
[11] ProctorExam. (2021). Infrastructure for Online Proctoring & Invigilation.[Online]. Available: https://proctorexam.com/
[12] Respondus. (2021). Assessment Tools for Learning Services. [Online]. Available: https://web.respondus.com/ Available:https://www.mdpi.com/1660-4601/18/1/282
[13] H. Hung and D. Gatica-Perez. Estimating cohesion in small groups using audio-visual nonverbal behavior. IEEE Trans. Multimedia, 12(6):563–575, 2010.
[14] X. Li, K.-m. Chang, Y. Yuan, and A. Hauptmann. Massive open online proctor: Protecting the credibility of moocs certificates. In ACM CSCW, pages 1129–1137. ACM, 2015.
[15] X. Liu. Video-based face model fitting using adaptive active appearance model. Image and Vision Computing, 28(7):1162–1172, July 2010.
[16] X. Liu, N. Krahnstoever, T. Yu, and P. Tu. What are customers looking at? In Proc. IEEE Conf. Advanced Video and Signal Based Surveillance (AVSS), pages 405–410, London, UK, Sept. 2007.
[17] X. Liu, T. Yu, T. Sebastian, and P. Tu. Boosted deformable model for human body alignment. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pages 1–8, Anchorage, Alaska, June 2008. IEEE.
[18] L. Nguyen, D. Frauendorfer, M. Mast, and D. Gatica-Perez. Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior. IEEE Trans. Multimedia, 16(4):1018–1031, 2014.
[19] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. Int. J. Computer. Vision, 42(3):145–175, 2001.
[20] M. Reale, S. Canavan, L. Yin, K. Hu, and T. Hung. A multi-gesture interaction system using a 3-d iris disk model for gaze estimation and an active appearance model for 3-d hand pointing. IEEE Trans. Multimedia, 13(3):474–486, 2011.
[21] W. Rosen and M. Carr. An autonomous articulating desktop robot for proctoring remote online examinations. In Frontiers in Education Conf., 2013 IEEE, pages 1935–1939, 2013.
[22] J. Roth, X. Liu, and D. Metaxas. On continuous user authentication via typing behavior. IEEE Trans. Image Process., 10:4611–4624, Oct. 2014.