Image-Hashing-Based Anomaly Detection for Privacy-Preserving Online Proctoring

Niharika Gurram, Rohith Reddy YarramReddy, Susmitha Jejigari, VenkataSai Vorsu

Huiyuan Yang

Abstract:

Online testing has become synonymous with online examination and proctoring has become an important element of it. However, many of the current solutions involves using video surveillance where third-party services conduct the examinations. This carries a lot of implications for student's privacy because their environment as well as data are disclosed. For these problems, we introduce the following privacy-preserving online proctoring system based on image hash function. Our system also targets the courses for peculiar movements like over face or body movements when attempting to cheat during an exam. Notably, the proposed system can work even if the student's face is occluded or obscured through blurring or a mask in the frame of the video sequences while still performing well. In this way, compared with raw videos, the system lowers the risk of privacy violations through analysing anonymized data. This paper adopted an in-house data set to test the performance of the proposed system in identifying the anomalies and promoting fair exam practices. All these give this approach a giant strategy over other online proctoring methods as a scalable, secure, and privacy-conscious solution in complex education system.

Introduction:

What the problem is

The central issue is to make online examination transparent and non-prejudicial to the students' rights whilst protecting their right to privacy. Modern approaches to online proctoring raise different issues of violation of privacy, need for physical control, or the delegation of managing personal information to third parties. Such techniques alert students to potential privacy infringements; are costly and time consuming; and sometimes offer low reliability in cheating detection. Also, there is no dataset available that has details of students' exam-taking behaviour along with privacy preserving nature such that it will enable the researchers to develop and test effective techniques.

Why this problem is important

This problem is crucial as changes to on-line format are going to become more permanent for many schools and colleges, and tests play a significant role in evaluating students' abilities and knowledge. The accuracy of tests is vital to building public confidence in the academic outcomes. Also confined to students' privacy, this aspect needs protection to mitigate on ethical as well as legal implications. Hence, sustaining the balance between these two aspects of online learning is important in order to develop a safe yet secure learning platform.

The challenges faced in solving this problem

Balancing Precision and Privacy: A proctoring system should learn cheaters with excellence such that their privacy as learners is not breached. This should be accomplished through system design to achieve a balance between the amount of surveillance and reliability.

Computational Efficiency: Real time live video data feed is computationally expensive and identifying potential suspecting behaviour from the video stream also requires significant computation. Ideally, a scalable system should be capable of withstanding large volumes of student traffic and should not take time computing in proportion to the numbers.

Lack of Public Datasets: To date, there are very few datasets that link exam-taking behaviours, and privacy retention; such datasets that exist are not public hence proving a challenge when training and testing new systems.

False Positives: Most automatic detecting systems reveal many false alarms, labeling many natural behaviours as cheating, thus eroding credibility of the system.

Scalability: In their current highly centralized form, large scale proctoring systems are costly and not easily plausible in institutions with limited finances.

What existing solutions are available

Live Proctoring: Students are also prohibited from communicated and teachers supervise students in real-time through webcams.

Recorded Proctoring: They are taken and after that, the proctors look at the examination process on the camera for any concerning activity.

Automated Proctoring: One of the applications relies on machine learning techniques to process live or recorded video data in order to detect irregularity or cheating.

Why the existing methods do not work or their shortcomings

Privacy Invasion: Many approaches are designed to gather information in circumstances that demand a view of students' faces and the surrounding space, which raises privacy issues.

Data Exposure: It is found that many systems make their students' video data openly visible to third-party service providers, making the individuals at risk of misuse of their data or hacker attacks.

High Costs: Face-to-face proctoring and tape-recorded proctoring are very costly because of their heavy demand on manpower making them irrelevant for large-scale use.

False Positives and Reliability Issues: Automated check of cheating is also a problem as systems can either miss cases of cheating or report cheaters when there is none.

Limited Anonymization: Some systems sought to remove the faces from videos, which is not quite enough to protect privacy.

Computational Overhead: Most current approaches are computationally intensive thus not scalable in terms of cost or computational resources needed.

The solution which is proposed in this paper.

To design a privacy-preserving online proctoring system, this paper introduces an image-hashing-based anomaly detection. The system addresses the shortcomings of existing methods by:

Privacy Protection: Instead of capturing original video stream the system employs hashed or anonymized stream to identify such atypical actions as turning the sight to the side or other erratic movement.

Efficiency: The approach is efficient in terms of computation which then translates into low resource consumption necessary in live video analysis.

Scalability: Automated detection dispenses with the constant need to supervise large populations of students, thus making the large-scale administration of deterrence cheaper.

Reliability: Instead of observing for visual abnormalities, the system correlates possible abnormal behaviours; thus enhancing its ability to flag exact oddities without getting overly sensitive and flagging unrelated actions.

Ethical Design: Being ethical in its approach, the system maintains academic standards while protecting individual's rights and prevents cheating while upholding a good ethical sense.

This is a novel approach to the means of a modern, elastic, and privacy-preserving proctoring an online examination, which can effectively eliminate the key problems of existing solutions.

Proposed system applications

1. Online Education and Exams

The software solution can be incorporated into Learning Management Systems and serves as a handy tool for Universities, Schools, and Certification programs to guarantee equitableness and safety of the tests and quizzes. It addresses the question of privacy, as well as academic dishonesty, therefore suiting remote examinations.

2. Human Resources Management and Training and Certification Programs for Corporate.

It can be added that organisations which organise and administer remote training and certification tests can employ the system for knowledge validation while adhering to privacy. It guarantees accurate results at the same time as it does not include anyone's personal details.

3. Government and Large-Scale examinations or tests are well-known examinations all around the world Answered by Teketora Lwelula.

The system is appropriate to use with large groups, for instance civil service, entrance exams, or any other standard base exams. Due to its ability to scale, it is practically cost-free and complies with regulations that governments and institutions of thousands of candidates must adhere to.

4. Professional Licensing Exams

This system will be of benefit for industries that need licensing exams for example; healthcare, law, and finance as it supports fairness and security on the exams. Privacy aspects are particularly meaningful in the case of professionals' information protection.

Proposed System

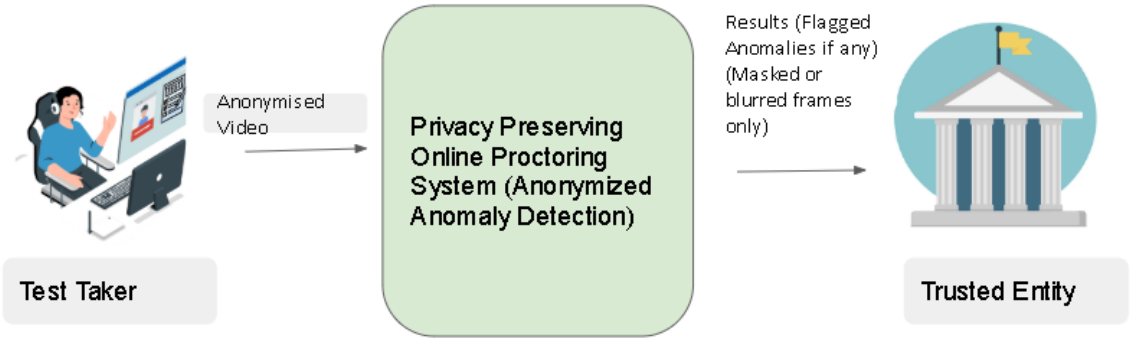

Figure 8: Proposed System

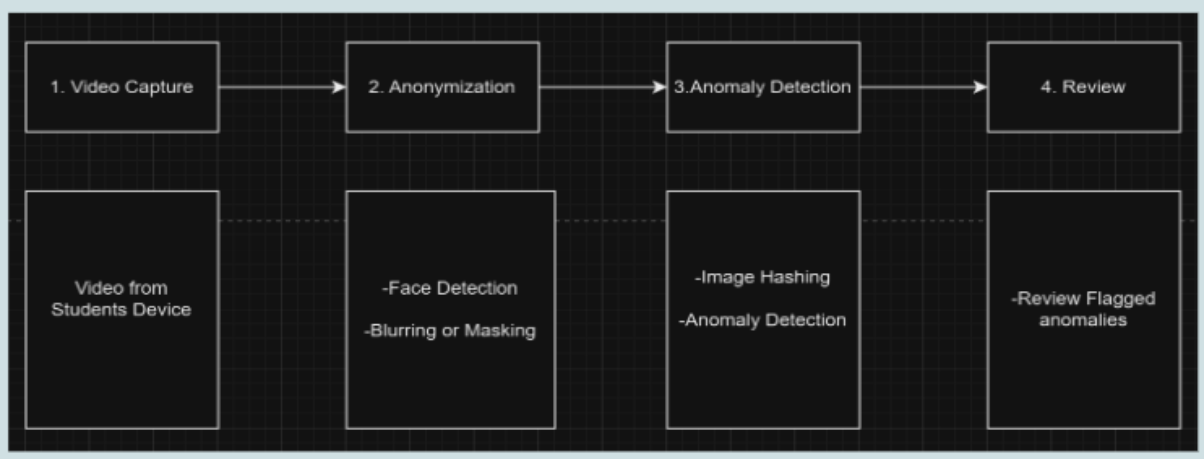

Figure 9: Components of Proposed System

Overview of the proposed method

The general approach to privacy-preserving online proctoring with anonymized anomaly detection is systematically developed providing the required protection of students privacy as well as the cloaking of online exams' integrity Figure 8. The process is broken down into four main steps: recording, blurring, outlying observation, and surveillance. Below is a detailed explanation of each step:

Dataset

In-House Dataset: One of the challenges we faced during this project was lack of a video dataset that would be available to the public and solely involves exam-taking and cheating-attempt

situations. For this purpose, we prepared an in-house dataset including four videos captured from four participants who are 24-26 years old. Every video is approximately 20 to 30 seconds based on the simulated exam activities and 4-10 seconds of calibration video. Cheating scenarios were mimicked by the means of introducing various anomalies as shown in the Figure 1. During normal examination circumstances it is anticipated that a student would mainly look at the computer terminal or the keyboard. All other movements or gestures involving the upper body are considered as possible abnormalities. Several positions indicated in figure 1. show the positions that students may assume during an exam. While "Front Face" is a normal behaviour of the students who are writing exams, other poses are like cheating where a student quickly looks at a cheat sheet or flips through a book.

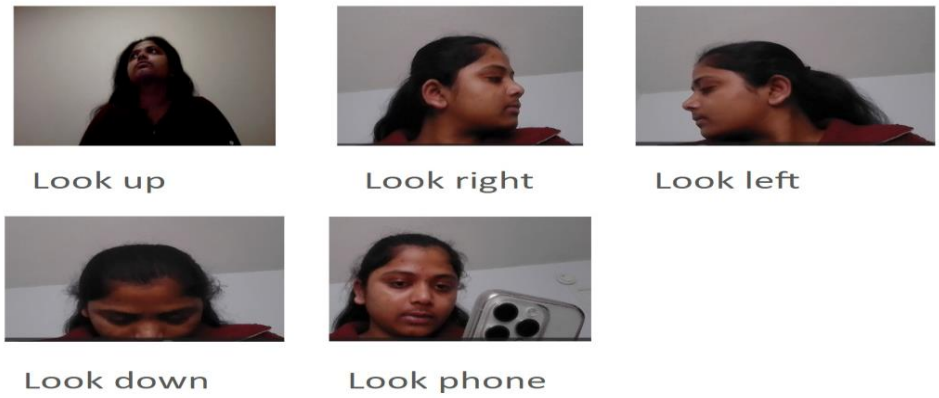

Figure 1: Various Types of Anomalies

Public Datasets: Because of the smaller size of the in-house dataset, we additionally confirmed the results presented in face and eye detection using the public image datasets including HELEN [1].

1.**The framework of the proposed method:**

The broad architecture composes of several elements which are compiled to enable privacy-preserving online proctoring. The test taker's video data is collected; the data are stripped of all personally identifiable materials; the data are scanned for behaviour consistent with cheating and the data are scanned for violations. It make it possible to input only such relevant data as is required and strip any input of identity tags in order to enhance the privacy of the examinees as well as protect the integrity of examinations as mentioned in Figure 9.

Step-by-Step Process:

1. Video Capture:

Initially student's video is captured before starting the exam which will be used as the calibration video. This Calibration video is used to compare the hash differences of the frames during the actual exam. It is also used to adjust the anchor frames dynamically during the exam.

Then the actual exam taking video is captured and stored. Both the videos (calibration and actual exam video) is very useful to the track the student's actions like eye movements, body

gesture indicating signs of cheating. The exam video is anonymised in the next step and then passed directly to the online proctoring system, which is then processed.

2. Anonymization (Face Hiding)

In addition, to remove privacy issues, the captured video is anonymised in the system. This is achieved through processes like face detection and subsequently covering the test takers face and or other aspects that identify him or her. The system also masks the face of the individuals captured in the video making sure that no one's identity is traceable, yet any form of suspicious activity is well detected. To achieve this, we first detected the face, eyes and then hidden the face using the bounding boxes Figure 2(a) and Figure 2(b) derived from the face and eyes detection module. These are our components in anonymization:

2**.1 Face and Eye Detection:** This step accomplishes the identification of the student's face and eyes in the video, where bounding boxes are drawn around detected face and eyes, which are later used for blurring or masking of facial attributes is used to preserve student's privacy.

In the proposed system, face and eye detection is a primary feature of the system that helps in blurring the faces of people involved in the video. This Face and Eye detection is done by using the machine learning where the given exam-taking video is analysed frame by frame. It segments bounding boxes around face and eyes, whose coordinates will later be used for masking or blurring the facial details. To achieve this, existing pre-trained machine learning algorithms are employed in this project including Dlib and MediaPipe and Haar Cascade, which boasts high detection rates as well as low resource utilization.

Face and Eye Detection using In-house Dataset

Thus, to assess the performance of the models, data experiments were carried out using an in-house dataset. The dataset was comprised of videos, in which face and eyes contours were drawn by human operator to use them as a ground truth. The models' performance was then compared based on the detection rate, calculated as:

Face Detection Rate = (Number of faces detected /Actual number of faces) × 100

Eye Detection Rate = (Number of frames on which eyes are correctly detected/ Number of frames labelled as face) * 100

This approach was useful to evaluate the efficiency of the face and the eye detection models used in the program. The results demonstrated that both detectors achieved favourable detection rates with MediaPipe being more effective given a higher FPS rate.

	Dlib				**Mediapipe**				**Haar cascade**			
	A	**B**	**C**	**D**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Face detection	81.98	87.93	55.87	100	99.67	100	100	100	80.50	79.86	30.98	100
Eye Detection	98.59	98.36	96.46	100	99.34	100	96.32	100	90.67	96.12	59.36	98.90

Table 1: Accuracy of Face and Eye Detection Using Different Libraries

(A: Niharika Gurram, B: Rohith Reddy YarramReddy, C: Venkata Sai Vorsu D: Susmitha Jejigari)

Facial and Eye Detection using HELEN Dataset

Subsequently, we conducted experiments on the HELEN dataset for face and eyes detection as well, because it contains images like the frames in the video. The dataset collected was critical to further test the model to determine its efficiency in face and eye detection in more varied situations. The Model was applied with the MediaPipe as it has a higher frames per second rate.

MediaPipe Model: This model identifies both face and eyes separately; the rate of efficiency of this model in the identification of eyes is not the same as that of face.

	Helen Dataset (Mediapipe)
Face detection	97.41
Eye detection	100

Table 2: Accuracy of Face and Eye detection on public dataset

In both datasets, the found faces and eyes received bounding boxes as shown in Figure 2(a), Figure 2(b), and then moved to the further stages of the system.

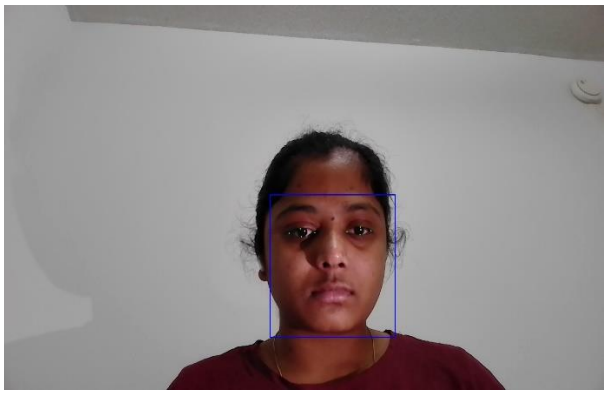

Figure 2 (a): Face and Eye bounding boxes

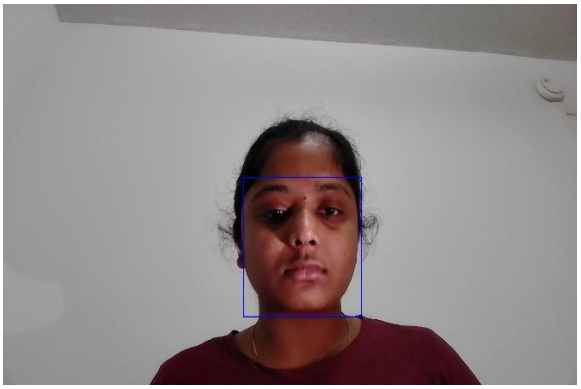

Figure 2(b): Face bounding boxes

2.2 Blurring or Masking:

In the current work, we employed bounding boxes obtained from preliminary face and eye detection for hiding face through blurring and masking of video frames. In particular, the face bounding boxes were used to define the areas to be hide, and the eye bounding boxes were used to avoid hiding of the eye regions. For blurring, Gaussian blurring was used with varying adjustable blur factor was applied only on the face area but without blurring the eye area as shown in the Figure 3(a). It also guarantees the optimal level of anonymization. On the other hand, for masking, binary mask (black and white) was generated to mask the facial area (except for the eyes) as shown in the Figure 3(b) using fixed bounding box. It is also computationally more efficient since the masking is used where black or white regions cover facial information to the extent that it becomes erased. These methods are flexible to be adaptively change and adjust so that certain masks or degrees of blurring can be delivered according to user's requirement. Together with the bounding box data, the described techniques allowed for accurate anonymization while excluding privacy-sensitive information, appropriate for additional analysis or dataset preparation.

Figure 3(a): Masking without eyes

Figure 3(b): Blurring without eyes

3. Anomaly Detection

The goal of the anomaly detection module is to detect suspicious activities in the anonimised video i.e., even if the student face is blurred or masked. To attain this, we exploit a technique for detecting anomalies in the exam viewing video using an image hashing based method.

A. **Image Hashing Method**

Perceptual hashing, also known as image hashing, is a way of computing hash values for two images given which we are able to compare them according to what hash values they have. Two images are said to be similar if the hash values of those two images are close to each other (difference below a given threshold). In this system we use image hashing to determine if all images in an exam-taking video are similar versus if we have various images that are quite different from one another indicating an anomaly.

Frame Similarity: If consecutive frames in the video are of the same shape it indicates the student maintained a position or the same pose. This is why if a frame stands out from the rest, you might think that the student is holding his/her pose differently and that might mean that the person is looking away from the screen or maybe making some weird movements.

Anchor Frame Comparison: In image hashing, each frame in the video is compared to a 'anchor frame' (referenced frame in the video) Hamming distance is used to calculate the difference of the frames hash values. A flag for potentially anomalous frame is set if the difference is above a predefined threshold.

To perform this task, we choose to use the popularity of the dHash (difference hash) method, which has been found to be a faster and more accurate image hashing than aHash (average hash) and pHash (perceptual hash) as mentioned in [2]. With a hash size of 12, the detection rate is high and the method is efficient.

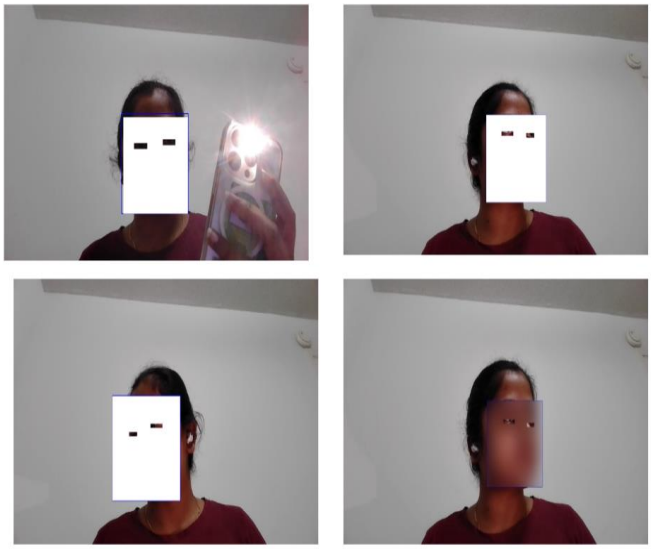

Figure 5: Different anomaly frames detected by model.

3.1. Determining the Threshold

A key step in detecting an anomaly is to set a threshold properly because the hash difference between frames can differ based on where the exam took place. The hash values can be influenced by the student's proximity to the camera or by the movements done during the exam.

As an example, if a student sits closer to the camera, the change in hash values between frames will be greater than the student sits farther away. Hash differences between the student's position and the machine's position vary, which makes determining a fixed threshold for detecting anomalies difficult. To deal with this the system uses a special threshold for each person exam session. The threshold of this threshold is dynamically calculated by the student's initial sitting pattern at the beginning of the exam. The process works as follows:

Initial Recording of Sitting Pattern (Calibration phase): The student's sitting position and eye interaction with the computer were recorded as the calibration video before the start of the exam. It takes the factors which include the student's posture and how he interacts on the screen. We assumed that this calibration video is not anomalous, and it is the ideal behaviour during the exam.

These calibration frames are then used to compute the initial threshold for comparison by computing the hash of one image and the hashes of all the other images. The one the maximum hash difference frame determined as the initial threshold. This calibration hash differences, and initial thresholds are stored to verify later with the hash differences of examination video obtained during the examination time. Calibration threshold is also considered later used to update the anchor frames, denoting that only major deviations are suspicious activities.

3.2. Selecting the anchor frame

We can determine the anchor frame image which is used to compare the hash difference between the images.

One approach is to use the first frame of the exam taking video as the fixed anchor frame and comparing all the other frames with it. This approach is not efficient because there may be slight changes in the body movement or sitting position during the exams which are not

anomalous but when the model compares it with the fixed anchor frame, due to hash difference it considers the frame as anomaly. For example, students can sit closer to the camera and away from the camera. When the student is closer the hash difference is low compared to anchor frame but during away position the hash differences are significantly higher. We address this issue by updating the anchor frame dynamically.

This issue of fixed frame is addressed by either adjusting the threshold or changing the anchor frame dynamically when the student position is changed. We discussed how to determine the threshold in the earlier section. To adjust the threshold will need student to record the calibration video again which is not user friendly. Instead, we applied dynamic change of anchor frames, where the threshold used during initial calibration is used during the entire exam.

The anchor frame can ideally be changed when the student change posture while sitting. This adjustment can be arrived at, by analysing the facial images captured in successive frames of the student. For example, when a student is closer the screen, his or her face will be wider in the frame and if the student is far from the screen, his or her face will be less wide in the frame. In this study, however, a different procedure is used to update the anchor frame to solve the problem. Rather than using punctual changes of the proximity level, the anchor frame is recalculated in normal behaviour in response to anomalous behaviour.

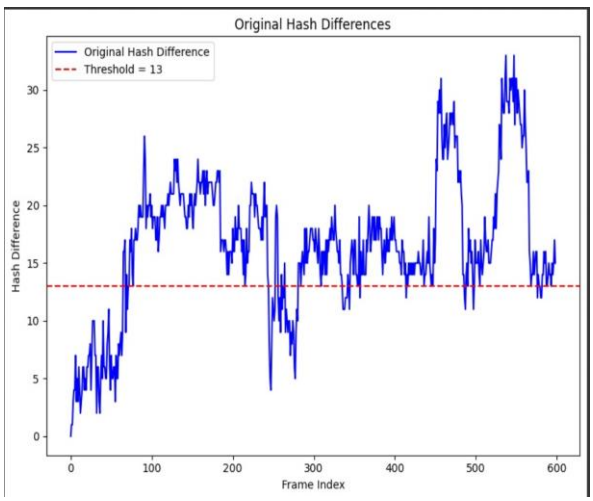

Figure 4(a): Hash Differences

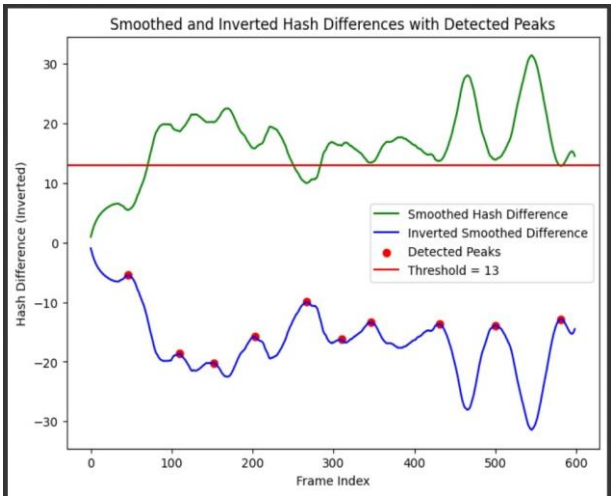

Figure 4(b): Smoothed and inverted hash differences.

Potential normal behaviours are depicted in the plot of the implicated behaviour shown in Figure 4(a) by points in the valleys of the curves. These valleys represent conditions where the student's attention is returned on the monitor, for instance after a brief look at a book (the downward points in the graph). Finally, before summarizing the findings, in signal analysis, utility is known and for removing the noise in the signal the Savitzky–Golay filter is employed. However, following the smoothing function the plot is reflected about the x-axis as shown in Fig 4(b). These frames that co-ordinate with the peak points of this inverted plot are marked in red are determined dynamically as new anchor frames

Thus, each time a new peak is observed (a red dot) the anchor frame is reset to the current frame and the hash differences are calculated from this anchor frame. It is shown in Figure 3(b) how this plot is done with the new anchor frames.

Results

An inhouse test dataset consisting of four exam taking videos recorded with different participants was used for the experiment. The videos, about 25 to 30 seconds long each, told participants to act as though in an exam environment, and try to cheat to create realistic situations from which to begin. The videos were manually labelled in each frame as either normal pose (normal exam behaviour), or an anomaly pose (suspicious activity such as looking away or excessive movement).

Participant	Accuracy	Recall	Precision	F1 Score
A	72.79%	97.73%	86%	0.79
B	81.87%	81.12%	95.87%	0.88
C	84.04%	86.18%	81.66%	0.84
D	81.41%	33.80%	86.75%	0.49

Table 3: Performance Metrics with privacy preserving methods (Mask)

Face and eye detection is performed using Media Pipe, Dlib and Haar Cascade models, followed by Gaussian blurring and single white masking to fully anonymous the detected faces. Media pipe has performed efficiently for face and eye detection on inhouse and public dataset from the results of Table 1 and Table 2. dHashing based image hashing of hash size 12 was performed to detect anomalies (and consequently detect significant differences in subsequent frames).

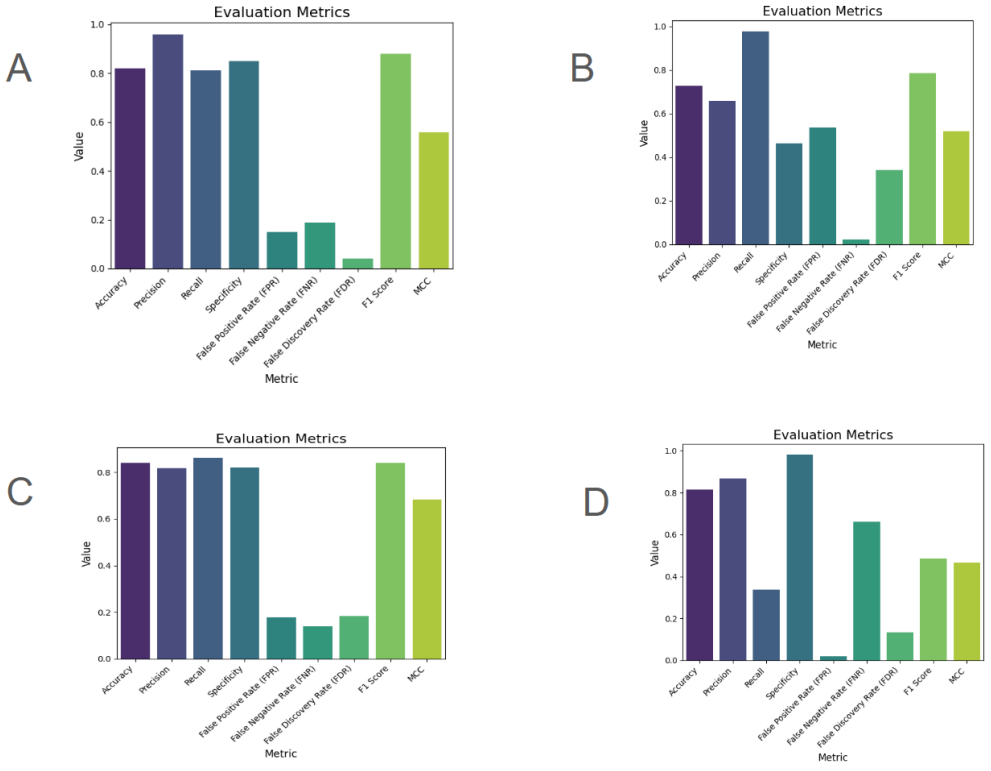

Figure 6: Evaluation Metrics for Anomaly detection for 4 participants (A,B,C,D)

Anonymisation is applied on the detected bounding boxes (Fig 2(A) and Fig 2(B)) as shown in figure 3(a) and Figure 3(b). Face and eye detection, anonymisation and hashing modules were

sequentially run on each video frame. MediaPipe has performed efficiently with high accuracy for Face and Eye detection, hence we used MediaPipe in blurring, masking and also anomaly detection. The accuracy and effectiveness (Table 3) of the proposed system was compared with the ground truth labels and finally the detected anomalies examples as shown in Figure 5. were compared with the detection accuracy. The evaluation metrics were calculated and plotted as shown in Figure 6. The anomaly detection model performed better in higher video quality and visibility.

Conclusion and Future work:

In this paper, we proposed a privacy-preserving online proctoring system using image-hashing-based anomaly detection. The experimental results have shown very promising outcomes. There are several ways this work can be further improved.

First and foremost, the important step is to build a complete dataset of students taking exams for robust analysis. Second, the proposed approach could be further improved by exploring more privacy-preserving techniques and a wide range of anomalies.

Contribution:

Time	Tasks	Personnel	Comments
Nov 1 - Nov 7	**Project Setup & Initial Data Collection**	Susmitha, Rohith	Define system requirements, finalize datasets (in-house and public), set up development environment for anonymization and hashing components. Ensure all necessary libraries and tools for face anonymization and image hashing are installed and tested.
Nov 8 - Nov 14	**Anonymization Implementation & Testing**	Niharika, Susmitha	Develop the face-blurring/masking component, validate anonymization effectiveness using HELEN dataset.conduct peer reviews to assess anonymization.

Nov 15 - Nov 21	**Image Hashing & Anomaly Detection Model**	Rohith, Venkat	Implement dHash method, set up thresholds to detect significant frame-to-frame movement as anomalies. Experiment with various thresholds and compare to baseline movement patterns; consult team for optimal sensitivity.
Nov 22 - Nov 26	**System Integration & Initial Testing**	Susmitha, Niharika, Rohith	Integrate anonymization and anomaly detection modules; run tests for accuracy, privacy preservation, and efficiency metrics. Document initial testing results; analyze any false positives/negatives in anomaly detection, and adjust parameters as needed.
Nov 27 - Nov 30	**Model Calibration & Optimization**	Susmitha, Venkat	Refine anomaly detection thresholds, optimize frame processing rate to meet real-time requirements. Assess computational efficiency on test hardware; ensure that privacy-preserving features maintain minimal recognition rate.
Dec 1 – Dec 5	**Final Evaluation & Documentation**	Entire Team	Conduct comprehensive tests for accuracy, precision, recall, FPS; complete documentation and

			final report for submission. Compile findings, summarize privacy metrics, complete project report including challenges, results, and future work suggestions

References:

1. https://github.com/zhfe99/helen
2. Yaqub, W., Mohanty, M., & Suleiman, B. (2021). Image-Hashing-Based Anomaly Detection for Privacy-Preserving Online Proctoring. arXiv preprint arXiv:2107.09373.
3. Nigam, A., Pasricha, R., Singh, T. and Churi, P., 2021. A systematic review on AI-based proctoring systems: Past, present and future. *Education and Information Technologies*, *26*(5),pp.6421-6445.
4. C. Cleophas, C. Hoennige, F. Meisel, and P. Meyer, "Who's cheating? mining patterns of collusion from text and events in online exams," Mining Patterns of Collusion from Text and Events in Online Exams (April 12, 2021), 2021.
5. J. Lemantara, M. D. Sunarto, B. Hariadi, T. Sagirani, and T. Amelia, "Prototype of online examination on molearn applications using text similarity to detect plagiarism," in 2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). IEEE, 2018, pp. 131–136.
6. S. Manoharan, "Cheat-resistant multiple-choice examinations using personalization," Computers & Education, vol. 130, pp. 139–151, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/ S036013151830304X
7. M. Bateson, D. Nettle, and G. Roberts, "Cues of being watched enhance cooperation in a real-world setting," Biology Letters, vol. 2, no. 3, pp. 412–414, 2006. [Online]. Available: https: //royalsocietypublishing.org/doi/abs/10.1098/rsbl.2006.0509
8. H. M. Alessio, N. Malay, K. Maurer, A. J. Bailer, and B. Rubin, "Examining the effect of proctoring on online test scores." Online Learning, vol. 21, no. 1, pp. 146–161, 2017.