WILEY | Hindawi

Research Article

Modeling and Classification of the Behavioral Patterns of Students Participating in Online Examination

B. J. Ferdosi ⓘ**, M. Rahman, A. M. Sakib, and T. Helaly** ⓘ

Department of CSE, University of Asia Pacific, Dhaka, Bangladesh

Correspondence should be addressed to B. J. Ferdosi; bjferdosi@uap-bd.edu

Online education has become an essential part of the modern education system, but keeping the integrity of the online examination remains a challenge. A significant increase in cheating in online examinations (from 29.9% before COVID-19 to 54.7% during COVID-19, as per a recent survey) points out the necessity of online exam proctoring systems. Traditionally, educational institutes utilize different questions in onsite exams: multiple-choice questions (MCQs), analytical questions, descriptive questions, etc. For online exams, form-based exams using MCQs are popular though in disciplines like math, engineering, architecture, art, or other courses, paper and pen tests are typical for proper assessment. In form-based exams, students' attention is toward display devices, and cheating behavior is identified as the deviation of head and eye gaze direction from the display device. In paper- and pen-based exams, students' main attention is on the answer script not on the device. Identifying cheating behavior in such exams is not a trivial task since complex body movements need to be observed to identify cheating. Previous research works focused on the deviation of the head and eyes from the screen which is more suited for form-based exams. Most of them are very resource-intensive; along with a webcam, they require additional hardware such as sensors, microphones, and security cameras. In this work, we propose an automated proctoring solution for paper- and pen-based online exams considering specific requirements of pen-and-paper exams. Our approach tracks head and eye orientations and lip movements in each frame and defines the movement as the change of orientation. We relate cheating with frequent coordinated movements of the head, eyes, and lips. We calculate a cheating score indicative of the frequency of movements. A case is marked as a cheating case if the cheating score is higher than the proctor-defined threshold (which may vary depending on the specific requirement of the discipline). The proposed system has five major parts: (1) identification and coordinate extraction of selected facial landmarks using MediaPipe; (2) orientation classification of the head, eye, and lips with K-NN classifier, based on the landmarks; (3) identification of abnormal movements; (4) calculation of a cheating score based on abnormal movement patterns; and (5) a visual representation of students' behavior to support the proctor for early intervention. Our system is robust since it observes the pattern of movement over a sequence of frames and considers the coordinated movement pattern of the head, eye, and lips rather than considering a single deviation as a cheating behavior which will minimize the false positive cases. Visualization of the student behavior is another strength of our system that enables the human proctor to take preventive measures rather than punishing the student for the final cheating score. We collected video data with the help of 16 student volunteers from the authors' university who participated in the two well-instructed mock exams: one with cheating and another without cheating. We achieved 100% accuracy in detecting noncheating cases and 87.5% accuracy for cheating cases when the threshold was set to 40.

1. Introduction

The fight for keeping the academic integrity of the examination can be traced back to the beginning of the education system. The COVID-19 pandemic accelerated online education though it was getting popular even before the pandemic [1].

As per the National Center for Education Statistics (NCES), USA, around 75% of undergraduate students took at least one online course in the fall of 2020 [2]. That shift to online learning has had a permanent effect on our education system even in the post-COVID-19 era. However, there are many challenges unique to online exams, including remote testing

issues, distrust of proctors, and concerns about cheating. According to a study by Newton and Essex [3], cheating in online exams jumped from 29.9% before COVID-19 to 54.7% during COVID-19. Another study during the COVID-19 pandemic found that 60% of students admitted to cheating on online exams [4]. Online exams offer more opportunities for cheating compared to traditional ones due to their remote nature and reduced surveillance. The availability of digital devices makes cheating tempting and harder to detect, often resulting in higher grades for cheaters.

General recommendations made for preventing cheating in online exams are as follows: inform students about the academic integrity policy of the institution and the consequences of cheating, encourage students to obey the academic ethics and help them in increasing their internal control against cheating, use higher-order thinking question sets, etc. [5]. Higher-order thinking questions set up in online exams are widely recommended by academicians [6], but this type of question may be appropriate for higher-level classes but not, in general, all levels of courses, and it may induce even more desperate behavior and new techniques for cheating. In addition, according to a study of human behavior in light of psychology and behavioral theory of fraud, self-control and ethical knowledge usually do not work to prevent fraudulent acts [7]. Skinner suggested the scientific analysis of human behavior and actions so that the behavior can be altered even before committing any fraudulent activity [8]. It is necessary to have constant supervision and activity analysis to deter cheating opportunities and motivations.

Traditionally, educational institutes utilize different questions in onsite exams, such as multiple-choice questions (MCQs), analytical questions, and descriptive questions. However, for online exams, they mostly set MCQs using forms because of the availability of form-based proctoring systems to ensure integrity. But in disciplines like math, engineering, architecture, art, or other courses where paper and pen tests are typical, form-based exam proctoring systems will fail to identify the students' cheating behavior. Automated proctoring of different exam types should differ since students' normal and cheating behaviors differ in different exam types. In form-based exams, students' attention is toward display devices and deviation from that can be considered cheating behavior [9–14]. On the other hand, in the written exams, student's main attention is toward the answer scripts and the orientation of the head and eye may vary depending on individual body posture. During proctoring such exams, the system should be more tolerant of body movements since this type of exam will involve more movements. Therefore, special care needs to be taken while using students' body part orientation as indicative of cheating behavior. In proctoring such online exams, it will be logical if emphasis is given on the pattern and the frequency of the movements. There are different types of online proctoring systems in practice, such as Real-Time Manual Proctoring (RTMP) [15, 16], Real-Time Automated Proctoring (RTAP) [11, 17–22], and Recorded Automated Proctoring (RAP) [10, 23] systems. In RTMP systems, human proctors remain actively present online to identify cheating. Scheduling of the exams is required beforehand so that a dedicated proctor can be assigned for each student. Use of this type of system is not feasible when a large number of students need to take an exam at the same time. Getting a large number of competent proctors will be a challenge. In RTAP systems, proctoring is done in real time using artificial intelligence (AI), and human interventions (HI) are deployed to resolve cheating cases. The system can proctor exams for a large number of students efficiently. The RAP systems use AI and HI on recorded videos after the exam, not in real time.

The recent works on online exam proctoring systems use input from different types of devices such as fingerprint readers, eye tribe trackers, wear cams, webcams, microphones, touch devices, mouse, and keyboards. Expensive devices like fingerprint readers or eye trackers are not available to everyone as these devices do not come as a package with a computer. The use of these devices with the proctoring systems will impose a financial burden on the students. The use of wear cams in the system will create physical discomfort for the students since these devices need to be worn on the student's heads. There are different formats of online examination, but not all formats have extensive use of mouse and keyboards, if the exam is not such that where answers are needed to be given through mouse and keyboard. The use of input from multiple devices will create processing overhead which will prevent the system from being real time. The online exam proctoring system is the demand of the rapidly expanding online education system. However, Lee and Fanguy in [6] rightly indicated that such a proctoring system should not come at the expense of students' comfort, additional financial burden, or additional anxiety.

Different behavioral data, such as head orientation, eye orientation, and voice, are used in online proctoring systems. Several systems use these as low-level features and transform them into high-level features to classify cheating behavior. Such systems work as a black box and refrain us from understanding the cheating behavior in terms of their actual behavior. That reduces the transparency of the proctoring process. Several other systems individually report head movement or eye movement to the human proctor and do not propose any unified model of cheating behavior. It is natural for humans to have movements of different body parts, and individual orientations of these body parts cannot directly confirm inappropriate behavior. To understand the student's cheating behavior, a unified model is required which will analyze the combined behavior of the relevant body parts, especially the head, eye, and lip.

In this work, we proposed an RTAP type proctoring solution for paper- and pen-based online exams using input only from a webcam. Our approach tracks head and eye orientations and lip movements in each frame and defines the movement as the change of orientation. Our system has five major parts:

(1) Identification and coordinate extraction of selected facial landmarks using MediaPipe

(2) Orientation classification of the head, eye, and lips with K-NN classifier, using coordinates of the landmarks

(3) A penalty system for change of orientation in a sequence of frames and identification of abnormal movement if the number of changes of orientation exceeds the proctor-defined threshold

(4) A cheating score calculator for each student at the end of the exam that penalizes abnormal movements, more abnormal movements higher than the cheating score, and a cheating case identification if the score exceeds proctor-defined threshold which can be adjusted depending on the exam scenario to be less or more tolerant to movements

(5) A visual representation of the movements of each student and cheating score graphs of a group of students to assist the human proctor in taking appropriate preventive action

The main contributions of this research are as follows:

(i) An online proctoring system for pen- and paper-based exams that considers specific requirements for such exams. None of the previous works consider the specific requirements for proctoring online pen- and paper-based exams. Previous systems are for form-based MCQs, which limit in-depth assessment

(ii) A very tolerant proctoring solution for the students. It also allows proctors to adapt the system as required. The previous works implement online proctoring at the cost of students' comfort and penalize them even because of their natural movements

(iii) A system that facilitates both cheating detection (by calculating cheating scores) and cheating prevention (by visualization of the behavioral pattern)

(iv) A system that will enable academic institutes to have traditional pen-and-paper exams in the online environment without compromising integrity

The rest of the paper is organized as follows: Section 2 describes the overview and basic definitions related to online examinations and proctoring systems, related research works are discussed in Section 3, the methodology of the proposed system is stated in Section 4, details of the datasets used can be found in Section 4.1, discussion on the system performance is in Section 5, and we conclude the paper in Section 7.

2. Overview of Online Examination and Proctoring

2.1. Online Examination. An online examination is an examination where students are allowed to sit for the academic assessment process at a location of the student's choice using one or more devices through the Internet. Students and academic proctors are separated by location but connected by digital devices and the Internet. Different formats of offline exams are in use and can be adapted to online examinations. Here, we discuss the formats of the online exams in

terms of the types of questions used to assess the student and the way of answering the questions, as follows:

(i) Type 1: exam with objective-type questions and answers given fully using digital devices

(ii) Type 2: exam with a combination of objective-type questions and descriptive questions; and answers are given either entirely using digital devices or partially with pen and paper for the descriptive part

(iii) Type 3: exams with descriptive questions and answers are given using pen and paper

2.1.1. Type 1. In this format, students are assessed with multiple choice, multiple choices with multiple answers, fill in the blanks, true/false, etc., type questions. It evaluates students objectively, and the assessment result can be generated immediately after the exam. Extensive use of a computer screen, mouse, and keyboard is needed. Thus, students' faces and gaze of eyes will be directed toward the screen most of the time during the exam period as seen in Figure 1(a). In this case, the placement of the device is closer to the examinee; thus, better view of the examinee is possible. During online proctoring of this type of examination detecting the position of the face, eye gaze using a camera, movement, and use pattern of mouse and keyboard will be helpful to identify cheating behavior.

2.1.2. Type 2. In this format, some parts of the examination questions are set using multiple choice and true/false-type questions, and some parts can be short answer, problem-solving, or essay-type questions. The part with objective-type questions can be assessed automatically while the part with descriptive-type questions needs to be assessed by the examiner. Therefore, the complete results of this type of exam can not be published immediately. The device usage of this type of exam is similar to type 1 exams if students need to answer both objective and descriptive questions using a mouse and keyboard. Thus, a similar approach to type 1 can be used to identify cheating. However, if the descriptive part of the exam is answered using pen and paper, then the body orientation will not face the computer screen directly while answering that part, eye gaze will be lowered and less visible, and there will be less or no mouse movement or keyboard usage (as shown in Figure 1(b)). Identifying cheating in this type of exam will be challenging.

2.1.3. Type 3. There are several courses for which assessment using objective-type questions is not enough to judge the depth of knowledge and high-level thinking questions are required. It is sometimes also preferred by the institutions to keep the traditional written pen- and paper-based exam conducted online where students need to submit the scanned copy or image of the answer script after the exam. Later, the examiner needs to check the scripts manually. Usually, this type of exam is administered to prevent cheating by using high-level thinking questions; it can create more stress and induce cheating behavior in students. Thus, they may try to get help from different online and offline materials and collaborate with
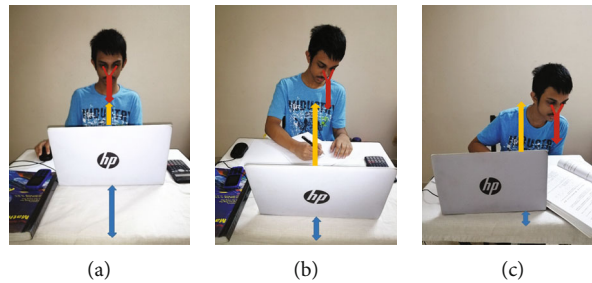
FIGURE 1: (a) Type 1 exam. (b) Type 3 exam. (c) Cheating behavior in type 3 exam.

friends or other acquaintances. As seen in Figure 1(b), in this type of exam, head orientation and eye gaze will deviate from the screen direction. This deviation is normal for this type of exam. In type 1, this deviation can be an indicator of cheating. Additionally, the students need to keep their answer script in front of the device; thus, the distance of the student from the camera will increase (indicated with a blue line in Figure 1). Because of the distance and different body orientations, identifying facial features is a challenging task.

2.2. Academic Dishonesty aka Cheating. Academic dishonesty or cheating usually refers to activities that violate the institution's academic integrity policy [24]. Any dishonesty in an academic examination, assignment, or any other assessment process such as the use of prohibited materials and copying from such materials, communicating or collaborating with others to get answers, unauthorized use of technology, and having someone else sit on exams on behalf of oneself can be considered as academic cheating. Any exam type discussed in the previous section is susceptible to cheating, but the deviation of body part movement will have a different pattern depending on the type of cheating which can be the indicator of cheating. If a student is copying or taking help from unauthorized material, both head posture and eye gaze will deviate substantially from normal (depending on the type of exam) posture since the student will keep the material in such a position so that it is not visible in the camera. If a student talks with another person in the room or on a communicating device, then there will be eye and lip movements. However, special care needs to be taken while declaring the students' body part movements as indicative of cheating behavior since the movement of body parts is natural human behavior and varies from person to person. In proctoring pen-and-paper online exams, it will be logical to give emphasis on the pattern and the frequency of the movements.

2.2.1. Types of Cheating. In offline exam halls, students are restricted to be confined in a particular space and prohibited from carrying or keeping unauthorized materials with them, and proctors remain vigilant to identify any unusual behavior or any unauthorized materials. On the other hand, in the online exam, students are not confined to any single space, and imposing restrictions on them is very difficult. Burgason et al. indicated in their study that "students have a detailed understanding of multiple ways to cheat" in an online exam [25]. As per their study, students use multiple resources for cheating such as opening multiple web browsers or multiple tabs in the browser, using materials written in Word documents, and using other devices like smartphones. There are many other ways of cheating that can be summarized as follows:

(i) Someone else is taking exams on behalf of a registered student

(ii) Use offline materials such as existing notes or books

(iii) Use offline materials on the test-taking device

(iv) Use the Internet in the test-taking device using multiple web browsers or multiple tabs of the web browser

(v) Use smartphones, smartwatches, or another computer

(vi) Collaborate with others using digital devices and communication platforms

(vii) Collaborate with others using an earpiece via Bluetooth

(viii) Receive help from other people on the test-taking premises

(ix) Keeping cheat material in a different place and using them

(x) Leaving the view of the surveillance camera using still images of one's own during video monitoring so that students can remain outside of the surveillance

2.2.2. Cheating Prevention Techniques. There are several recommendations in the research literature [24, 25] and also in the current practices of educational institutions to prevent cheating behavior such as the following:

(i) Educating students about the academic integrity policies of the institution and the consequences of cheating

(ii) The zero-tolerance attitude of the institution toward cheating

(iii) Making cheating difficult by using

(a) time-intensive exams

(b) randomized questions and randomized answers

(c) one question at a time with no backtracking and limited time to answer

(d) locked down browsers

(e) open-ended questions

(f) using assessment methods other than exams

(g) online exam proctoring tools to increase the chances of getting caught thus at least creating the psychological effect in reducing cheating

2.2.3. Online Exam Proctoring. Traditionally, proctoring an exam is the process of invigilation of the test-taking environment to ensure academic integrity so that there is no suspicion in the assessment. The worldwide revolution of online education has opened the gate to quality education for everyone irrespective of their geological position or financial condition. It also raises questions about the quality and integrity of online education, especially stems from difficulties in keeping the integrity and authenticity of the remote examination process. These questions about online examinations created the demand for the online proctoring system. This type of proctoring system is an essential tool for online exams. It helps educators to identify and prevent any unethical behavior during the online exam. Over the years, researchers proposed different types of online proctoring systems. These systems can be grouped into two broad categories: Real-Time Exam Proctoring and Recorded Exam Proctoring. These two groups can be categorized further into two subgroups: Manual Systems and AI-Based Automated Systems, as shown in Figure 2.

Real-Time Exam Proctoring. In this type of system, proctoring is done remotely through the online examination system using video and/or audio from web cameras or inputs from other devices, when students are taking their exams.

Recorded Exam Proctoring. In this mode, recorded audio, video, or other required data of the students' examination session is used to identify the cheating behavior after the exam is completed.

Manual Systems. A real-time or recorded exam can be proctored by a human being as done in the traditional offline exam hall proctoring, but in online exams, it is done remotely. It is required that a human proctor remains available and vigilant during the whole extent of the online exam. The human proctor may or may not take help from machine intelligence, but full-time human labor is mandatory. Extensive dependency on human proctors can restrict the number of students sitting for exams at a time in real-time mode or may require a longer time to declare an exam as cheat-free in recorded mode because of the lack of expert proctors.

Artificial Intelligence-Based (AI-Based) Automated Systems. In this type of system, the primary judgment about the exam integrity is done using artificial intelligence and, later on, seeking intervention from human proctors if required. Some automated systems halt the examination process if any instances of unethical behavior are identified until the decision is finalized by a human proctor. Some other systems report possible cases of cheating to the human proctor, and the final verdict is from the human proctor. In such systems, human proctors do not need to be vigilant during the entire exam, and the time of the experts is saved. It also allows educational institutes to administer the exam for a large number of examinees together.

3. Related Works

With rapid digitalization, the world is experiencing changes in every aspect. The education sector is no different. After the start of online education in 1985 [26], a large number of people are receiving education through this new technology. To compete with the traditional systems, the online system has to overcome several challenges. Assessing students' academic performance remotely is one of the most difficult challenges. The research on online exam proctoring tools can be traced back to the beginning of the 21st century [6]. There are a vast number of publications that can be found on online exam proctoring in recent years, but in our work, we considered the publications that are published in ISI or Scopus-indexed journals or conferences. A lot of work on offline exam hall proctoring is also available, but we have limited our discussion only to the work on online exam proctoring. We have discussed the related research on online exam proctoring based on the categorization stated in Section 2.2.3.

3.1. Real-Time Manual Proctoring (RTMP) Systems. In this type of proctoring, human proctors remain vigilant throughout the exam period in real time. The Live+ solution of ProctorU [15], a commercial proctoring system, provides RTMP-type services where their dedicated expert human proctors proctor and control the exam in real time. The system allows proctors to take assistance from machine intelligence during their proctoring task. Educational Testing Service (ETS) uses ProctorU for conducting at-home testing [16] of TOEFL and GRE.

3.2. Real-Time Automated Proctoring (RTAP) Systems. The system proposed by Li et al. combines both automatic and manual approaches to detect cheating [20]. The automated system flags probable cheating behavior, and ambiguous cases are further investigated by peer students and then by the authorities. The system has two modules, a webcam-only system and a multimodal detection system consisting of a webcam, a gaze tracker, and an EEG sensor, but the authors argue for the use of multiple sensors. The system identifies any suspicious activity and assigns a flag for such activity. A total number of flags assigned will determine the cheating and noncheating behavior.

Atoum et al. proposed an automated process for exam proctoring [11]. They presented a multimedia analytic system that performs automatic online exam proctoring. They used a webcam, a wear cam, and a microphone to monitor the visual and acoustic environment of the testing location. The system had six basic components to detect cheating behavior: user verification, text detection, voice detection, active window detection, gaze estimation, and phone
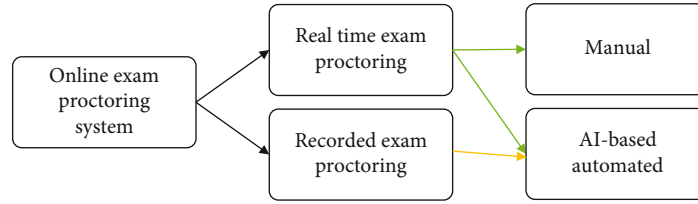
FIGURE 2: Category of online exam proctoring systems.

detection. By combining these estimations and applying a temporal sliding window, they have designed high-level features to classify whether the test taker is cheating.

Turani et al. compared the performance between a webcam and a 360-degree security camera in the existing proctoring systems [18]. The study was conducted based on video and sound recorded by the system and analyzed using machine learning algorithms. The proposed system used face recognition, eye orientation detection, and sound detection to detect cheating behaviors. If any student moves their face more than 30 degrees or makes any attempt to talk, an alert will pop up on the screen, and receiving three alerts will lead the system to close the exam. In other cases, the system flags and records the video if it detects any unusual behavior to be later checked by a proctor, reducing the need for real-time proctoring and the number of proctors.

Garg et al. proposed an approach to supervising online exams using the webcam [19]. The system detects faces using Viola-Jones and recognizes faces using the Haar cascade classifier and CNN. The system first detects the face and tags them according to their name, which they will use during the sign-up process. The model is then trained with the tag and the face. The face detection module detects if the face of the examinee is there. If no face is detected, then it is considered cheating, and the exam is terminated. The system tracks if the face of the candidate has moved out of the screen, as well as the number of faces on the screen. If it is more than two, then it is considered cheating.

Li et al. proposed a visual analytic approach to assist a human proctor in determining whether a student is cheating [17]. They constructed a visual representation of student behavior by analyzing the webcam footage of the students and the mouse movement during the exam. The system identifies and visualizes the head and mouse movements of the students, enabling course instructors and teachers to provide convenient and reliable proctoring for online exams. The system detects the head using faster R-CNN and identifies the head orientation using the pitch, yaw, and roll of the head. The system also computes a risk factor for a student in a particular question.

Ganidisastra and Bandung proposed a fully automated exam proctoring system where they used two methods: face detection and face identification [21]. If the system cannot find a face or the user's face does not match the face of the registered examinee, the system terminates the exam. Photos for training are taken during the student's attendance in the classroom.

The system proposed by Jia and He identifies cheating behavior using audio-visual data. The system ensures the identity of the student with face recognition and then ana-

lyzes facial expression, eye, and mouth movement and any audio data if any to identify the cheating behavior [22].

3.3. Recorded Automated Proctoring (RAP) Systems. Chuang et al., in their exam proctoring method, have placed particular emphasis on two factors: head attitude variation and time spent responding to questions after the question has been asked [10]. The video of the exam was recorded via a webcam, and the time spent on each question was also recorded. The recorded video was analyzed after the exam session, and the data were extracted based on the visual focus of the examinee. The extracted data were then differentiated into six different head-pose measurements (position and rotation of the X, Y, and Z axis) and eight statistical features. The final model was constructed of a hierarchical logistic regression combined with time delay. One of the limitations of the study stated by the authors was the false alarm rate of 0.102 on average, meaning one out of 10 instances was falsely identified as cheating leaving a chance of false identification in a large number of instances.

Masud et al. proposed an automated cheating detection technique that classifies student behaviors into cheating and noncheating using four features: head movement, eye movement, mouth opening, and examinee's identity extracted from the recorded video [23]. They considered the video as a multivariate time series data and used CNN, BiGRU, and both RNN and LSTM together, for the time series data. They also tested two traditional algorithms: random forest and logistic regression for the task.

From the study of the related works, we have the following observations:

(i) Data from different devices such as webcams, 360-degree security cameras, gaze trackers, EEG sensors, speakers, mouse, and keyboards have been utilized for identifying cheating. It is not financially feasible for the students or for the academic institute to use special devices like 360-degree security cameras, gaze trackers, and EEG sensors for proctoring. Additionally, the use of multiple devices will make the proctoring system computationally expensive and prevent it from performing in real time

(ii) The involvement of human proctors is inevitable to have a final verdict of cheating behavior, but to increase the scalability of the online proctoring systems, taking help from machine intelligence is also inevitable

(iii) Movements of the body parts such as the head, eye, and lip can play an important role in identifying

cheat behavior. In recent research, these body part orientations are considered as an indicator of cheating. However, a particular orientation should not be considered cheating especially in pen-and-paper exams. A person can have a particular orientation while writing on paper which may deviate from the screen direction. It is more important to observe the change of orientation, which indicates a movement

(iv) Several systems penalize students even for their natural movements. Students can move their heads as a part of natural body movement, gazing around while thinking, and some students mutter while writing answers. However, the frequency of natural movements and cheating movements differs. An efficient proctoring system should be able to identify this difference

4. Methodology

There is an idiom saying "Actions speak louder than words" which can be considered as a universal truth for understanding human behavior. Thus, it can also be very effective for identifying cheating in online exams where students are equipped with new techniques for cheating and supported by new technologies. It is sometimes very hard to identify materials and devices used for cheating, but it is also very hard to avoid basic human behavior while cheating. Human activity modeling for understanding behavioral patterns has been getting the attention of the research community recently [27]. Such models consider psychological and physiological explanations of why and how human body parts move and coordinate with each other for conducting a task. In neuroscience research, the "vision for action" refers to visual processing in the brain required for performing any activity, such as reading a book or catching a baseball. In "vision for action," the guided and coordinated movement of the body parts is required [28, 29]. Research also shows that while reading, a person usually moves his/her head toward the reading material first, and then, eye gaze moves in the direction of the material [30]. Therefore, if students want to copy from the material available in their environment, then both head and eye movements will be present. There is also a motor-memory trade-off in coordinated movements of the eyes, head, and body in a copying task [31]. The study shows that if the cheating material is nearby, students will rely less on their memory and look frequently toward the cheating material and there will be more head-eye coordinated movements. As in vision for action, research shows that there are coordinated movements of the head and lips while speaking as well [32, 33]. People most commonly nod to show agreement, acceptance, or acknowledgment and shake their heads from side to side to show disagreement, etc. These spontaneous non-verbal-coordinated movements will be helpful in identifying cheating.

Considering these psychological and physiological facts, we have proposed a system identifying the head, eye, and lip movements from real-time video. A penalty model is formulated in the system where the coordinated actions of

TABLE 1: Landmark coordinate datasets.

Datasets	No. of landmarks (each with x, y, and z coordinates)	Columns	Rows
Head orientation dataset	5	15	48,828
Eye orientation dataset	22	66	65,769
Lip orientation dataset	10	30	84,278

head-eye-lip or head-eye or head-lip, etc., are penalized more to get a cheating score. Online exam proctoring systems generally have two parts: one is authentication, and the other is cheating detection. In our work, we concentrate on cheating detection and we intend to identify cheating behavior based solely on the face, eye, and lip movements. In our system, we have the following assumptions:

(i) Exams are administered with descriptive questions, and answers are given using pen and paper (type 3 exams as described in Section 2.1.3). Hence, students will concentrate most of their time writing on paper, not on the computer screen

(ii) Video input from only one web camera, thus no additional cost on devices for the students

(iii) The view of the camera mostly covers the facial regions

(iv) Minimal usage of mouse and keyboard

(v) Students are smart enough to hide or keep the cheating material out of the camera view

(vi) The microphone can be muted or not available

(vii) If students are out of the camera view or hide their faces, it will be considered cheating behavior

4.1. Data Collection. In our work, we used two types of datasets: the first one is the video datasets, and the second one is the landmark coordinate datasets collected from the first. We collected the video dataset from 16 student volunteers from the author's university, University of Asia Pacific (UAP), Dhaka, Bangladesh, where 9 are males and 7 are females. Among the participants, 4 were graduate students, and 12 were undergraduate students. All of the participants were between the ages of 20 and 30 years old. We took consent from all volunteers for using and publishing their videos and ethical approval from the Research Ethics Committee (REC), UAP. An exam environment was set before the start of the exams, and students were provided with questionnaires as well as cheating materials. Students were given two sets of questions, one before recording and another during the recording of the session. Volunteers prepared with one of the questions that were provided before the exam session. One session was recorded with the question that students were prepared to answer, and another session was recorded with a new question, and students were provided with cheating materials. Volunteers were also instructed to
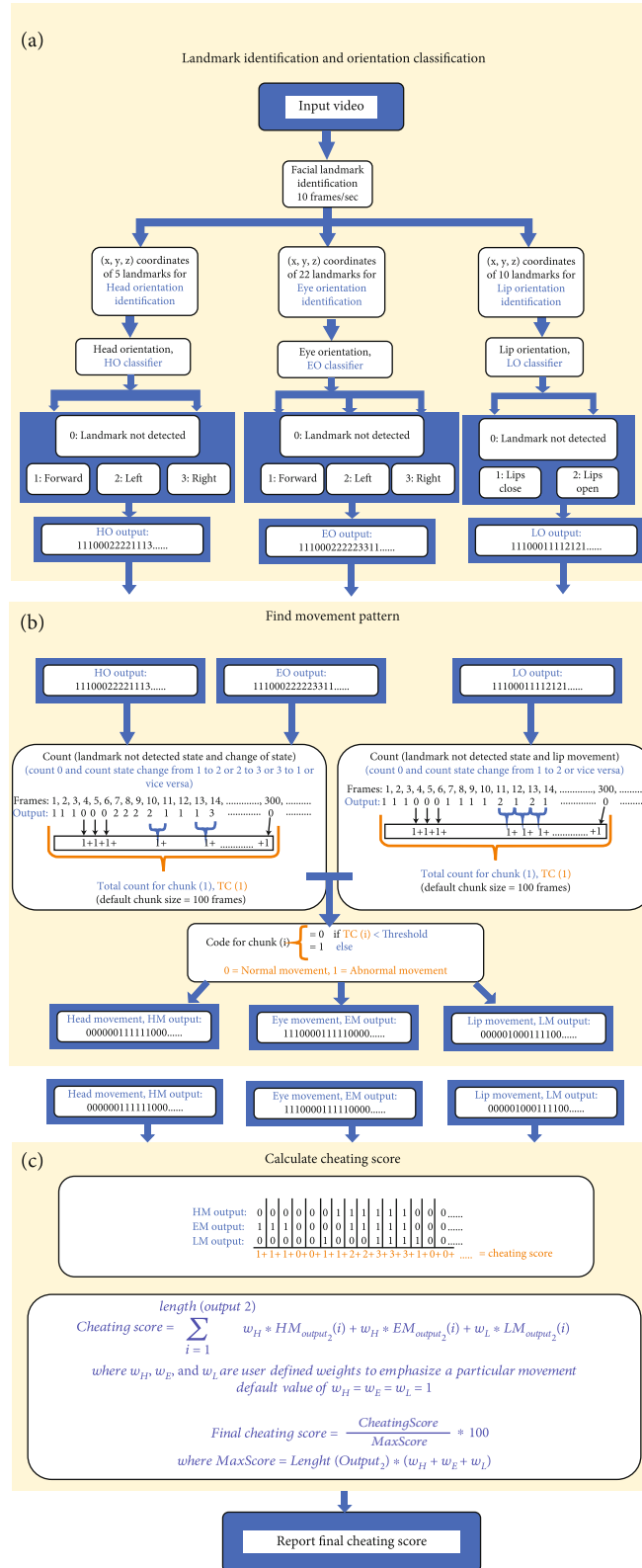
FIGURE 3: Basic working principle: (a) landmark identification systems; (b) finding movement pattern; (c) calculating cheating score.

do some of the cheating that frequently happens during an online exam, such as browsing the phone, talking with friends over the phone, talking with someone in the room outside the webcam view, and looking at cheating materials from a different angle. The video datasets contain all the recorded sessions, and the landmark coordinate datasets contain the extracted landmark coordinates from the videos with a 30 frames/sec rate.

Input: $Video\ Input$, V_{input}
Output:
Coordinates of Landmarks for Head Orientation, HOL_xyz,
Coordinates of Landmarks for Eye Orientation, EOL_xyz,
Coordinates of Landmarks for Lip Orientation, LOL_xyz
1: $Frame \longleftarrow GetFrame(V_{input})$
2: $FaceMesh \longleftarrow MediapipeFaceMesh()$
3: **if** $(FaceMesh == True)$ **then**
4:    $[LM_{idx}, LM_{coord}] \longleftarrow GetLandMarks(FaceMesh(Frame))$
5:    **for** i = 0 to 477 **do**
6:      **if** $(LM_{idx}[i] == [1,33,61,263,291])$ **then**
7:        $HOL\_xyz \longleftarrow LM_{coord}[i]$ ▷ Find Coordinates of Landmarks for Head Orientation
8:      **end if**
9:      **if** $(LM_{idx}[i] == [13,14,81,82,87,178,311,312,317,402])$ **then**
10:        $LOL\_xyz \longleftarrow LM_{coord}[i]$ ▷ Find Coordinates of Landmarks for Lip Orientation
11:      **end if**
12:      **if**
13:        $((LM_{idx}[i] \geq 468$ **and** $LM_{idx}[i] \leq 477)$ **or**
14:        $LM_{idx}[i] == [33,133,144,153,158,160,263,373,380,362,385,387])$ **then**
15:        $EOL\_xyz \longleftarrow LM_{coord}[i]$ ▷Find Coordinates of Landmarks for Eye Orientation
16:      **end if**
17:    **end for**
18: **else**   ▷No face detected, thus no landmark coordinates
19:    $HOL\_xyz \longleftarrow null$
20:    $EOL\_xyz \longleftarrow null$
21:    $LOL\_xyz \longleftarrow null$
22: **end if**

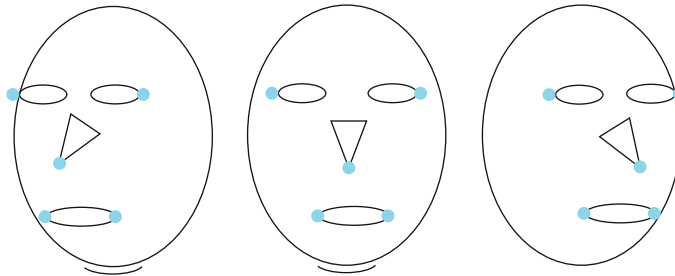ALGORITHM 1: Find coordinates of the landmarks.

FIGURE 4: Change of prominent landmarks during left and right orientation of the head.

4.1.1. Training Datasets. To train the models to identify the orientation of the head, eye, and lips, we prepare a training dataset from the collected videos from the mock exams.

(1) Video Datasets. We have separated the dataset according to the major movements of the head, eye, and lips that have an impact on making a predictive behavior model for cheating detection in online exams. Keeping that in mind, we have separated the video dataset based on three different orientations, i.e., head orientation, eye orientation, and lip orientation.

   (i) Head orientation video data: head movement videos have been separated into "LeftOrientation," "RightOrientation," and "ForwardOrientation." Each section contains 13 videos, and each video is 90 seconds long on average. In these videos, the stu-

dents were moving their heads from the normal position to either left, right, or forward. There were two more sections of videos where only the left movements of the students and only the right movements of the students were cut and combined into one shorter video clip with an average length of 40 seconds. According to the movements of the head, the extracted landmark data points were labeled to create the final dataset using those points

   (ii) Eye orientation video data: eye movement videos are also categorized into three different orientations as described in Section 4.1.1. Eye orientations based on the orientation of the iris are separated from the exam scenario videos. These videos contain eye movements with real scenarios. The eye movements are coordinated head-eye movements with partially
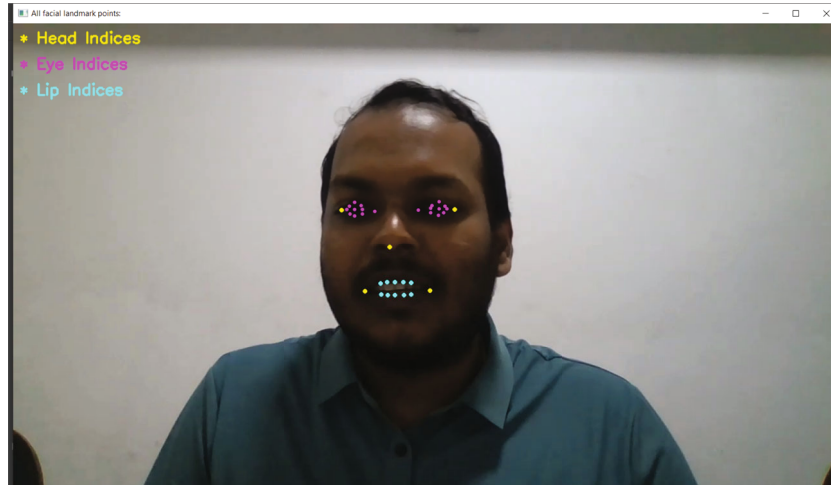
FIGURE 5: Landmarks used: yellow dots represent landmarks used for identifying head orientation; magenta dots represent landmarks used for identifying eye orientation; cyan dots represent landmarks used for identifying lip orientation. *Some of the yellow and magenta dots overlap each other.

visible iris. We also created videos with different orientations where the iris is visible

(iii) Lip orientation video data: for the lip orientation datasets, we took two types of videos. For the "LipOpen" class, we took video segments while students were talking with someone or over the phone. For the "LipClose" class, we took video segments where students were not talking

(2) Landmark Coordinate Datasets. From the video datasets, we get facial landmark coordinate datasets. Facial landmarks are the important points on one's face that can be helpful to identify a person or their emotion or attitude. Lugaresi et al. developed an open-source framework named MediaPipe for building a perception pipeline that facilitates finding facial landmarks [34]. We identify the landmarks on the video dataset using their framework and utilize selective landmarks for our work. Here, we generated three datasets for head, eye, and lip orientations containing the x, y, and z coordinates of those landmarks. In Table 1, the summary of the three datasets is shown.

4.2. Cheating Detection. The working principle of the proposed system has four major steps:

(i) Step 1: facial landmark identification and landmark coordinate extraction (Figure 3(a))

(ii) Step 2: orientation classification (Figure 3(a))

(iii) Step 3: find the movement pattern (Figure 3(b))

(iv) Step 4: cheating score calculation (Figure 3(a))

4.2.1. Facial Landmark Identification and Coordinate Extraction. MediaPipe creates a complete face mask and 3D coordinates of the 478 facial landmarks and other outcomes for face and facial expression detection. Our proposed

algorithm (Algorithm 1) for landmark identification and coordinate extraction selects five landmarks for the head, ten for the lip, and twenty-two for the eyes and extract coordinates of those landmarks.

The eyes, nose, and lips are the most prominent regions of a face. Especially, if the head moves left and right, then substantial changes that can be identified in the front camera view are the two corners of the eyes, the nose tip, and the two corners of the lips (see Figure 4). Hence, we used these five landmark indices (1, 33, 61, 263, and 291) for detecting head orientation. Among these five landmarks, two are on the outer corners of the two eyes, one on the tip of the nose, and two on the corners of the lips (indicated with yellow color in Figure 5). To detect lip orientation, we use ten landmark indices: 13, 14, 81, 82, 87, 178, 311, 312, 317, and 402. Out of these ten indices, five are on the upper lip and the rest on the lower lip (indicated with cyan color in Figure 5). For eye orientation detection, twenty-two landmark indices are used [35]. Indices 468 to 477 are landmarks on each eye comprised of one on the center of the eye pupil and four on the outer iris circle (indicated with magenta color in Figure 5). We use 12 more eyelid landmarks indices: 33, 133, 144, 153, 158, 160, 263, 373, 380, 362, and 387.

After identification of the selected landmarks, our algorithm extracts the x, y, and z coordinates of the landmarks. If the system is unable to detect any faces due to the absence of the student from the camera view, there will be no landmarks, and a null value is assigned to the coordinates. In later stages, it will be used as an indicator of absence from the camera view and will be considered cheating behavior. The coordinates of the landmarks constitute the datasets to train the models for movement classification and input to the trained model to classify the students' movements.

4.3. Orientation Classification. Our proposed methodology for orientation classification is stated in Algorithm 2 and also in Figure 3(a). The head is the largest visible body part among the three, and its movement is more distinguishable

Input: *HOL_xyz, EOL_xyz, LOL_xyz*   ▷Output from Algorithm 1
Output: *String* : *HOout_string[], EOout_string[], LOout_string[]*
   ▷ **Head Orientation Classification**
1: **if** (*HOL_xyz == null*) **then**
2:    *HOout_string[]* ⟵ 0   ▷ **No face detected**
3: **else**
4:    **if** (*HOClassifier(HOL_xyz) ==* "*ForwardOrientation*") **then**
5:      *HOout_string[]* ⟵ 1
6:    **else**
7:     **if** (*HOClassifier(HOL_xyz) ==* "*LeftOrientation*") **then**
8:      *HOout_string[]* ⟵ 2
9:     **else**
10:     **if** (*HOClassifier(HOL_xyz) ==* "*RightOrientation*") **then**
11:       *HOout_string[]* ⟵ 3
12:     **end if**
13:    **end if**
14: **end if**
15: **end if**
   ▷ **Eye Orientation Classification**
16: **if** (*EOL_xyz == null*) **then**
17: *EOout_string[]* ⟵ 0   ▷ **No face detected**
18: **else**
19: **if** (*EOClassifier(EOL_xyz) ==* "*ForwardOrientation*") **then**
20:    *EOout_string[]* ⟵ 1
21: **else**
22:   **if** (*EOClassifier(EOL_xyz) ==* "*LeftOrientation*") **then**
23:    *EOout_string[]* ⟵ 2
24:   **else**
25:    **if** (*EOClassifier(EOL_xyz) ==* "*RightOrientation*") **then**
26:     *EOout_string[]* ⟵ 3
27:    **end if**
28:   **end if**
29: **end if**
30: **end if**
   ▷ **Lip Orientation Classification**
31: **if** (*LOL_xyz == null*) **then**
32: *LOout_string[]* ⟵ 0   ▷ **No face detected**
33: **else**
34: **if** (*LOClassifier(LOL_xyz) ==* "*LipsClose*") **then**
35:    *HOout_string[]* ⟵ 1
36: **else**
37:   **if** (*LOClassifier(LOL_xyz) ==* "*LipsOpen*") **then**
38:    *LOout_string[]* ⟵ 2
39:   **end if**
40: **end if**
41: **end if**

ALGORITHM 2: Orientation classification.

than the others. Though we move our heads in a specific direction to look at something in that direction, it is possible to look the other way than the direction of the head, which is an important feature that makes us create different models for head movements and eye movements. The head movement and eye movement orientations are classified into "forward orientation," "left orientation," and "right orientation." We classify the lip orientations into "open" or "closed" to detect if a student is talking during the exam.

To classify the orientations of the head, eyes, and lips, we compared the performance of four machine learning models to identify the best model for the task: logistic regression [36], k-nearest neighbor (K-NN) [37], random forest [38], and support vector machine [39]. All four models are supervised machine learning algorithms and are suited for numerical data. Logistic regression performs better for categorical data with linear decision boundaries, K-NN is suited for data with nonlinear or complex decision boundaries, random

Input: $InString[]$: $HOout_{string}[], EOout_{string}[], LOout_{string}[]$ ▷ Output from Algorithm 2
Output: $OutBit$:
HeadMovementPattern: $HM_{BitString}[]$,
EyeMovementPattern: $EM_{BitString}[]$,
LipMovementPattern: $LM_{BitString}[]$ ▷ 0 or 1 where 0: Normal Movement and 1: Abnormal Movement
1: ChunkSize,$CS \longleftarrow 100\ frames$
2: TotalCount, $TC \longleftarrow 0$;
3: $Threshold \longleftarrow TH$ ▷ User defined value, default TH = 40
4: **for** i = 1 to CS **do**
5:     **if** $(InString[i] == 0)$ **then**
6:         $TC = TC + 1$
7:     **else**
8:         **if** $(InString[i] \neq InStrin[i+1])$ **then**
9:             $TC = TC + 1$
10:         **end if**
11:     **end if**
12: **end for**
13: **if** $TC \geq TH$ **then**
14:     $OutBit \longleftarrow 1$ ▷ 1: Abnormal Movement
15: **else**
16:     $OutBit \longleftarrow 0$ ▷ 0: Normal Movement
17: **end if**

ALGORITHM 3: Find movement pattern.

Input: ▷ Obtained from Algorithm 3
Head Movement Pattern: $HM_{BitString}[]$,
Eye Movement Pattern: $EM_{BitString}[]$,
Lip Movement Pattern: $LM_{BitString}[]$
Output: $CheatScore$ ▷Calculated cheating score for unusual movements
1: $ToatalScore \longleftarrow 0$
2: $w_H \longleftarrow 1 ; w_E \longleftarrow 1 ; w_L \longleftarrow 1$;   ▷User defined weights of the movements; default 1
3: $BitLength \longleftarrow Length(HMOut_{BitString})$
4: $MaxScore \longleftarrow BitLength * (w_H + w_E + w_L)$ ▷Maximum Score
5: **for** i = 1 to BitLength **do**
6:     $Score = w_H * HM_{BitString}[i] + w_E * LM_{BitString}[i] + w_L * EM_{BitString}[i]$
7:     $TotalScore + = Score$
8: **end for**
9: $CheatScore = TotalScore/MaxScore * 100$

ALGORITHM 4: Cheating score calculation.

forest performs better on data with nonlinear decision boundaries, and support vector machines can work both for linear and nonlinear class boundaries.

4.4. Find Movement Pattern. In our work, we define the movement of the body parts as their change of orientation. For example, if the head is oriented left and there is no change in the orientation, then we consider it as no movement occurred. If the orientation of the head changes from left to right or forward, then it is considered as a movement. Our proposed method for finding movement patterns is stated in Figure 3(b) and also in Algorithm 3. However, we can not detect a movement from a single frame as it requires a change in orientation rather than a specific one. Moreover,

to identify abnormal movements, we need to consider the movement pattern over a period of time rather than a single movement. For example, right-handed students can orient their heads to the left and can keep on writing without changing their posture for a period of time. This orientation or tilting of the head should not be considered cheating behavior. In our system, we consider two phenomena as indicators of cheating: (i) the student is out of the camera view, i.e., no face is detected, or (ii) frequent changes in the orientations of the head, eye, and lips. For this, we consider a chunk of 100 frames and count the number of times the student's face is not present in the camera view and the changes of orientations of the head, eyes, and lips (as described in Algorithm 3). If the total count is higher than
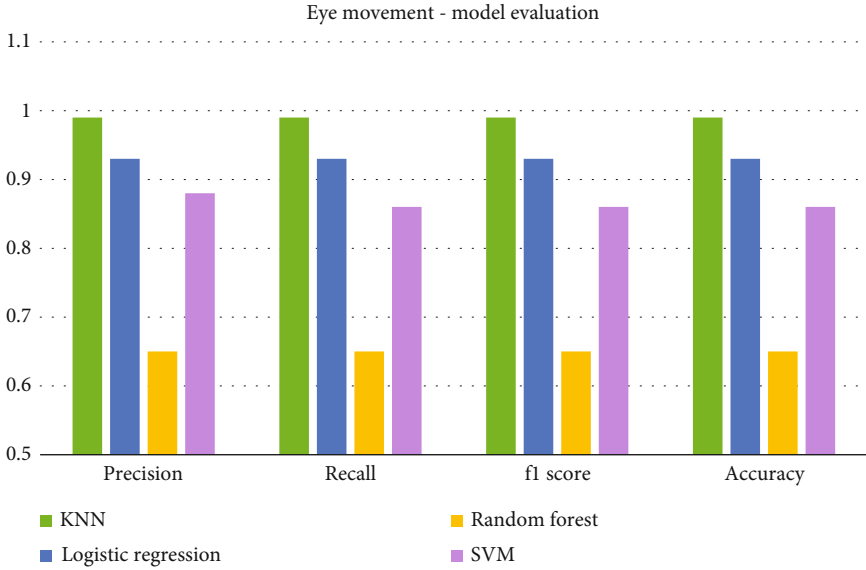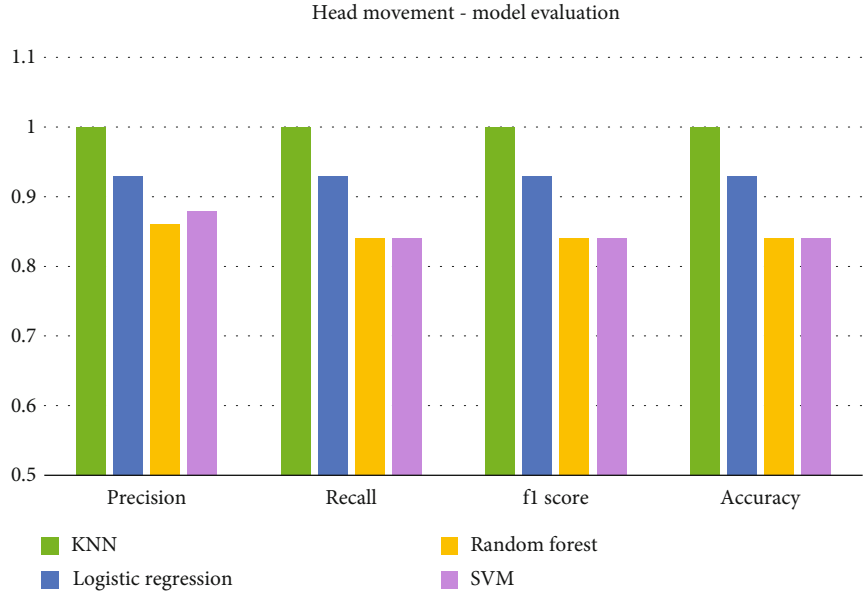
Head movement - model evaluation

(a)

Eye movement - model evaluation

(b)

FIGURE 6: Continued.

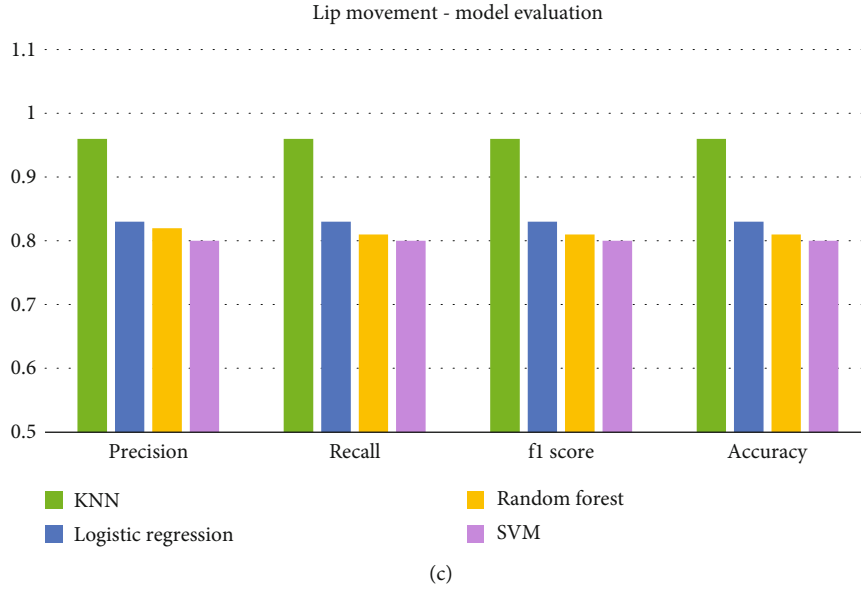Lip movement - model evaluation

(c)

FIGURE 6: Performance of the models for (a) head orientation dataset, (b) eye orientation dataset, and (c) lip orientation dataset.

a threshold value, then the movements are considered as an abnormal movement pattern. This process will prevent the system from reporting regular movements of the students as cheating.

4.5. Cheating Score Calculation. To identify the cheating cases, we calculate a cheating score as described in Algorithm 4 and Figure 3(c). Cheating score is the summation of the movement scores (0 for normal and 1 for abnormal movement) of the head, eyes, and lips for a period of time. If all three body parts have abnormal movements, then we will get the highest cheating score. If the human proctor (user of the proposed system) wants to penalize a particular body part's movements more than other body parts, it can be done using the weighting factors: w_H, w_E, and w_L. The final cheating score is calculated as the percentage of the maximum cheating score. We calculate the maximum cheating score by multiplying the summation of the weight variables with the length of the bitstring we get from Section 4.4. In our experiments, we set all the weights equal because we decided not to penalize any specific body part movement compared to others. From our observation of the students' behavior during exams, we noticed that some students have a natural tendency to read out loud what they are writing or thinking and some students shake their heads or move their eyes more frequently than others. If we weigh a specific body part more than the other, these students may have a higher chance of being penalized. Therefore, we set equal weights for the head, eye, and lip in our experiments. However, if any system user thinks otherwise, we keep the provision for penalizing any of the movements more than others by setting the weights differently.

5. Results and Discussion

Our goal for the study is to make a lightweight video data-based machine learning model that will help proctors detect
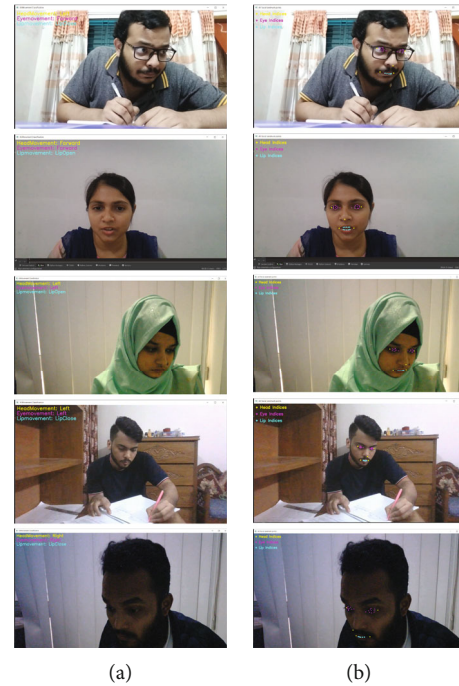

(a)   (b)

FIGURE 7: Orientation classification of the different examinees in different scenarios: (a) classification outcome; (b) position of the landmarks of that frame.

the cheating behavior of students in an online exam. Our target was to build a system that would require the students a PC or laptop with minimal specifications. We also kept in mind that we should not use any extra gadgets other than the webcam that is of minimal cost or comes built-in with many laptops. That is why we concentrate on identifying students' behavior with minimal data. We also use a minimum number of landmark coordinates from the videos so that the system can work in real time.
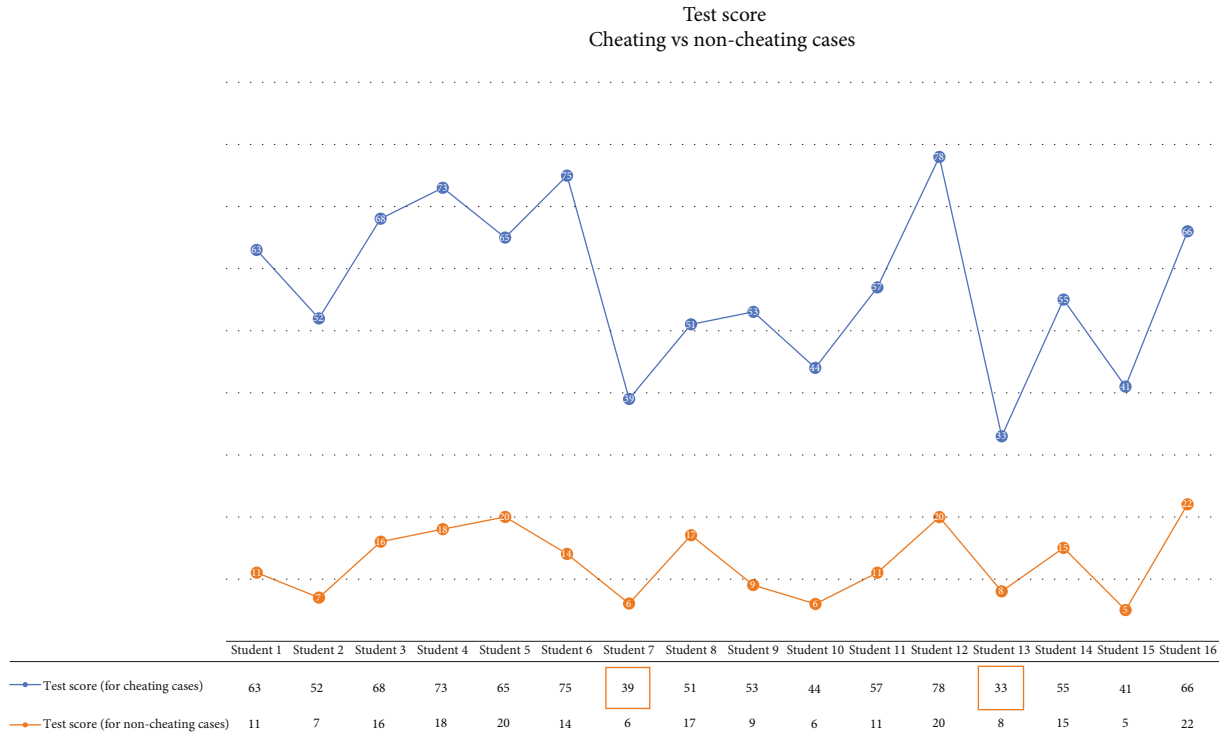
Test score
Cheating vs non-cheating cases

	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6	Student 7	Student 8	Student 9	Student 10	Student 11	Student 12	Student 13	Student 14	Student 15	Student 16
Test score (for cheating cases)	63	52	68	73	65	75	39	51	53	44	57	78	33	55	41	66
Test score (for non-cheating cases)	11	7	16	18	20	14	6	17	9	6	11	20	8	15	5	22

FIGURE 8: Cheating scores for cheating vs. noncheating cases.

5.1. Orientation Classification Accuracy. To classify the orientation of the head, eyes, and lips, we used different machine learning models and selected the model with the best performance for the task. We compare the performance of four different algorithms, K-NN, logistic regression, random forest, and SVM, in terms of accuracy, precision, recall, and $f1$-score. In head orientation classification, K-NN obtained the highest accuracy of 1.00, whereas logistic regression, random forest, and SVM obtained the accuracy of 0.93, 0.84, and 0.84, respectively. K-NN also outperforms other models in terms of precision, recall, and $f1$-score (see Figure 6(a)). K-NN also performs better in identifying the eye and lip orientations (see Figures 6(b) and 6(c)).

In Figure 7, snapshots of the examinations held in the different scenarios with different orientation patterns of the examinees can be observed. In Figure 7(a), classification of the students' head, eye, and lip orientations can be observed, and in Figure 7(b), positions of the landmarks are shown. Here, we can see students appearing in examinations with different backgrounds, lighting conditions, and different camera positions. Though there were strict instructions on how to sit before the camera, some were too close to the camera, even not facing it properly (row 5 of Figure 7). Some kept their camera a little too far than instructed (row 4 of Figure 7). Still, the system can observe their behavior.

5.2. Cheating Identification Accuracy. In our system, we calculate a score out of 100. The score is normalized by the maximum score possible, thus reflecting the frequency of the student's movements during the exam. Figure 8 shows the cheating scores of all 16 volunteers for both cheating and noncheating cases. The lowest and highest scores for the cheating cases are 33 and 78, respectively, whereas the lowest and the highest scores for the noncheating cases are 5 and 22, respectively. It is also apparent from the graph that in most cases, students who have a tendency to move frequently have relatively higher scores in both cases. A case will be marked as cheating if the cheating score is above a threshold value defined by the proctor. We set the threshold value to 40 in our experiment. We could set our threshold to the average of these two scores since there is a good margin between the highest score for noncheating cases and the lowest score for cheating cases. However, to make the system more robust, we use a threshold value of 40, higher than the minimum score of cheating cases. With this threshold, the cheating cases of students 7 and 13 remain as noncheating cases and thus unidentified in terms of score. Our system identifies 14 cheating cases out of 16 correctly, thus achieving an 87.5% success rate. For the noncheating cases, the system can identify all 16 of them correctly and achieve a 100% success rate. However, the strength of our system is that it can not only identify cheating but also prevent cheating by visualizing movement patterns so that a human proctor can intervene if excessive movements are observed.

6. Visualization of Students' Behavior

Prevention is always better than cure. Keeping this in mind, we proposed two visualizations in our system. One is for showing the movement pattern of the head, eyes, and lips. Another is for displaying the intermediate cheating score. A proctor can use these visualizations for early intervention.
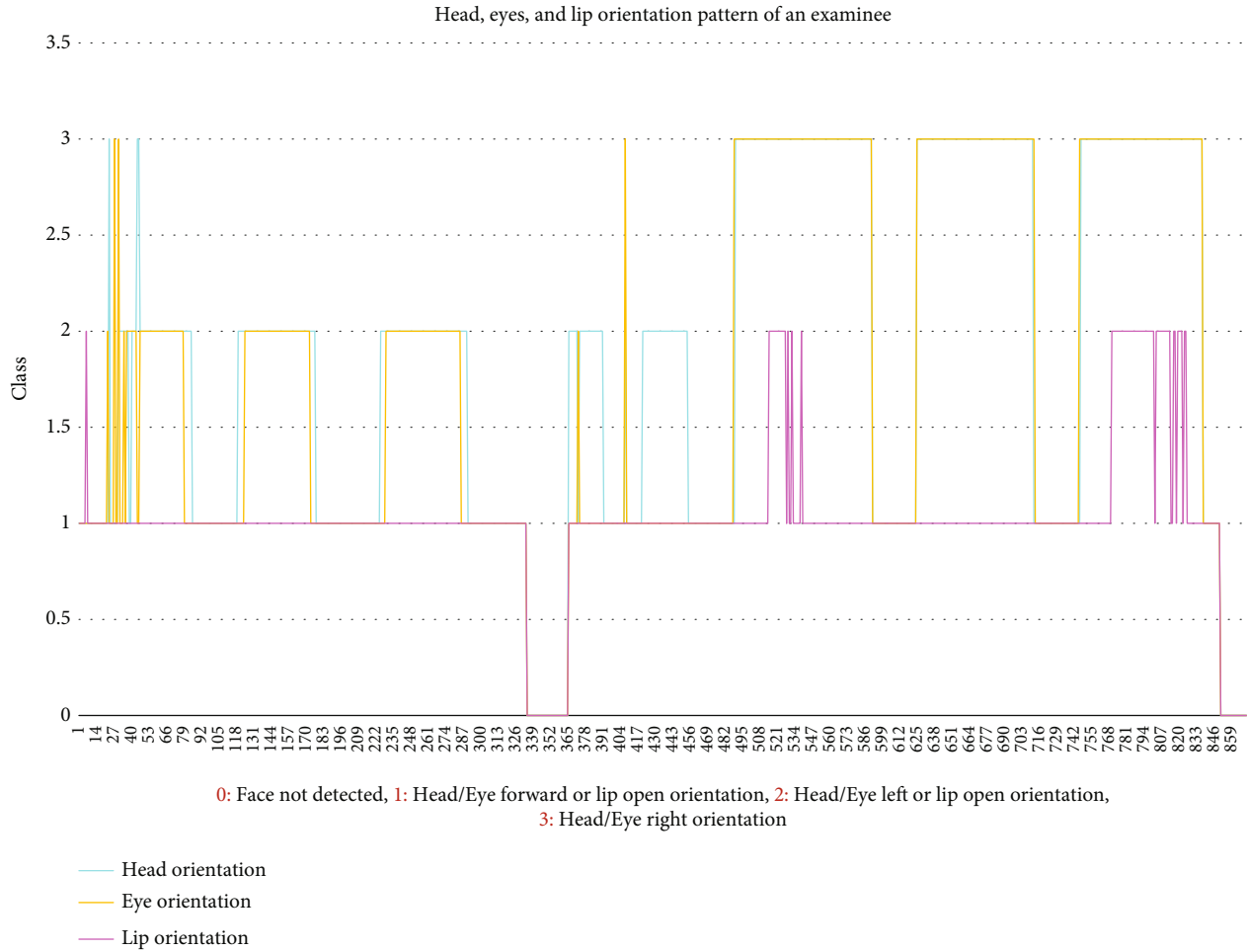
Head, eyes, and lip orientation pattern of an examinee

0: Face not detected, 1: Head/Eye forward or lip open orientation, 2: Head/Eye left or lip open orientation,
3: Head/Eye right orientation

——— Head orientation
——— Eye orientation
——— Lip orientation

FIGURE 9: Behavior pattern of an examinee.

6.1. Movement Pattern Visualization. In Figure 9, we represented an examinee's movement pattern in line graphs that an invigilator can observe during the exam period. Here, for the head and eyes, "forward," "left," and "right" orientations are represented with 1, 2, and 3, respectively, and for lips, "close" and "open" orientations are represented with 1 and 2, respectively (as discussed in Section 4.3). If the student is absent from the screen, then there will be no landmarks found; this phenomenon is represented with 0 for the head, eye, and lip orientations. From the graph, we can see that lines for the head and eyes overlapped each other in most of the cases which indicates their coordinated movements. Here, it can be observed that in earlier stages of the exam, the student's head and eye orientations were left or forward, but in the later stage, it was mostly right orientations, and there were several instances of lip movements too. There were instances of absence that can be observed in the graph.

6.2. Cheating Behavior Visualization. The system calculates a cheating score for every examinee based on their activity pattern. In Figure 10, a visual representation of the cheating or noncheating behavior of the participating students is depicted using multiple line graphs: one line graph per stu-

dent. In the graph, the left vertical axis shows the activity score of each student. In Section 4.4, the procedure of transforming frame-wise decisions into chunk-wise scoring is discussed, and the score can range from 0 to 3, where 0 means no head, eye, and lip movement and 3 means there are movements of all the body parts considered in the system. The right vertical axis shows the cheating score of the student calculated for 18 chunks, i.e., for the video of 1-minute time duration. The horizontal axis represents the chunks where each chunk consists of 100 frames. The color of the line graph indicates the level of cheating: the gray color indicates that the score is below the proctor (user of the system)-defined threshold and the red color indicates that the score is higher or equal to the set threshold and can be considered as potential cheating cases.

6.3. Limitations. Several aspects can be taken care of. In the current version, we have the following major limitations:

(i) No authentication module

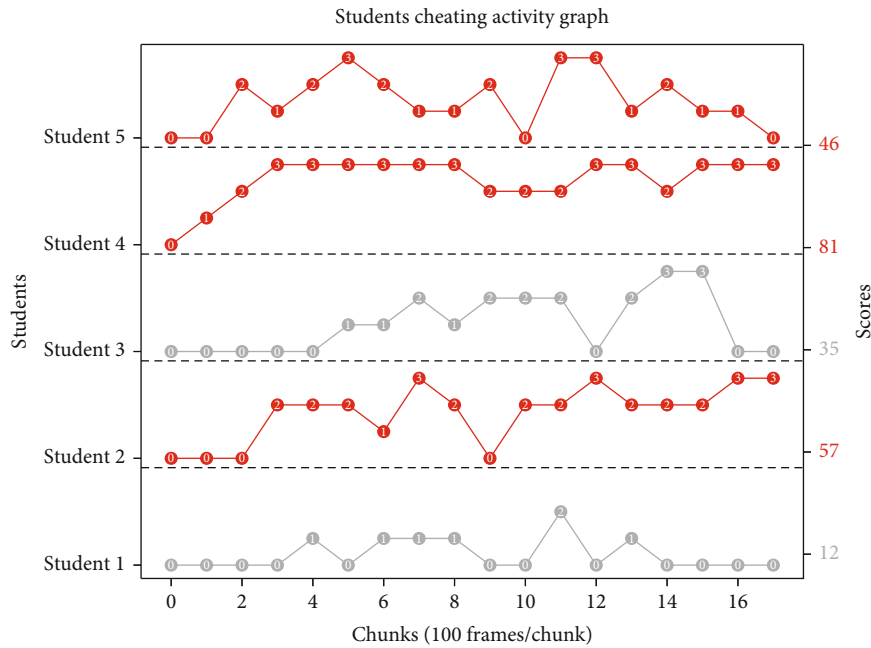(ii) Only works for videos from webcams of laptops or PCs and does not work for videos from mobile cameras

FIGURE 10: Students' cheating activity graph: gray lines are indicative of noncheating cases while red lines are indicative of cheating cases.

(iii) If examinees keep talking or looking around, hiding their lips or eyes using other objects, our system fails to identify such movements. MediaPipe saves the coordinates of the landmarks from the previous frames even if those landmarks are hidden in the current frame, and this is why our system fails to identify the change in those areas

(iv) False positive may occur if the natural tendency of the examinee is to move frequently (such as people with ADHD). False negatives may occur if the examinee cheats without much movement

(v) A formal user survey has not been conducted yet for the system

7. Conclusion and Future Work

In this article, we described a proctoring system for online pen- and paper-based examinations where we analyze the frame-by-frame orientations of the head, eye, and lips of the examinee, try to identify the pattern of movements of these body parts, calculate scores for the chunks (100 frames/chunk), and also calculate a final cheating score for the whole exam duration. We also report the movement patterns and cheating scores to the human proctor through visualization. We used the MediaPipe library to get the coordinates of the selected facial landmarks, and we used K-NN, the best-performing machine learning model for orientation classification. Analysis of the mock test participated by the volunteers confirms the strength of the proposed system. Our system obtained a success rate of 87.5% identifying cheating cases with proctor-defined cheating threshold set to 40.

In this research, we present an online proctoring system for pen- and paper-based exams. Our proctoring system is tolerant for the students, and proctors can adapt the system as required. The most important aspect of our system is that it facilitates both cheating detection (by calculating cheating scores) and cheating prevention (by visualization of the behavioral pattern). Our system will enable academic institutes to have traditional pen-and-paper exams in the online environment without compromising integrity.

In the future, we will try to overcome the identified limitations of the current version of the work. We will integrate the authentication module into our system to confirm that only registered students appear in the exam for the whole duration. We will also try to solve the issue of talking or looking around by hiding the lips or eyes. A formal user validation test will be organized to identify more issues to improve the current version.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

References

[1] S. Palvia, P. Aeron, P. Gupta et al., "Online education: world-wide status, challenges, trends, and implications," *Journal of Global Information Technology Management*, vol. 21, no. 4, pp. 233–241, 2018.

[2] Institute of Education Sciences US Department of Education, "Undergraduate enrollment. Condition of education," 2022, https://nces.ed.gov/programs/coe/indicator/cha.

[3] P. Newton and K. Essex, "How common is cheating in online exams and did it increase during the COVID-19 pandemic? A systematic review," *Journal of Academic Ethics*, vol. 8, 2023.

[4] A. A. Malik, M. Hassan, M. Rizwan, I. Mushtaque, T. A. Lak, and M. Hussain, "Impact of academic cheating and perceived online learning effectiveness on academic performance during the COVID-19 pandemic among Pakistani students," *Frontiers in Psychology*, vol. 14, article 1124095, 2023.

[5] M. Carla, "A study of cheating beliefs, engagement, and perception – the case of business and engineering students," *Journal of Academic Ethics*, vol. 17, no. 3, pp. 291–312, 2019.

[6] K. Lee and M. Fanguy, "Online exam proctoring technologies: educational innovation or deterioration?," *British Journal of Educational Technology*, vol. 53, no. 3, pp. 475–490, 2022.

[7] J. Wells, *Fraud Encyclopedia*, Association of Certified Fraud Examiners Publication, USA, 2005.

[8] B. F. Skinner, *Science and human behavior. Number 92904*, Simon and Schuster, 1965.

[9] R. Bawarith, A. Basuhail, A. Fattouh, and S. Gamalel-Din, "E-exam cheating detection system," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 4, 2017.

[10] C. Y. Chuang, S. D. Craig, and J. Femiani, "Detecting probable cheating during online assessments based on time delay and head pose," *Higher Education Research & Development*, vol. 36, no. 6, pp. 1123–1137, 2017.

[11] Y. Atoum, L. Chen, A. X. Liu, S. D. H. Hsu, and X. Liu, "Automated online exam proctoring," *IEEE Transactions on Multimedia*, vol. 19, no. 7, pp. 1609–1624, 2017.

[12] G. Fenu, M. Marras, and L. Boratto, "A multi-biometric system for continuous student authentication in e-learning platforms," *Pattern Recognition Letters*, vol. 113, pp. 83–92, 2018.

[13] A. Abozaid and A. Atia, "Multi-modal online exam cheating detection," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, pp. 1–6, Prague, Czech Republic, 2022.

[14] T. Potluri, S. Venkatramaphanikumar, and K. Venkata Krishna Kishore, "An automated online proctoring system using attentive-net to assess student mischievous behavior," *Multimedia Tools and Applications*, vol. 82, no. 20, pp. 30375–30404, 2023.

[15] "Meazure learning, proctoru, live+," 2023, Accessed on June 20, 2023, https://www.meazurelearning.com/exam-technology/proctoru-onlineproctoring/live-plus.

[16] "Ets - toefl, gre," 2023, Accessed on June 20, 2023, https://www.ets.org/gre/test-takers/general-test/register/at-home-testing.html\#accordion-1990b03ae5-item-80f1b77595.

[17] H. Li, M. Xu, H. W. YongWang, and H. Qu, "A visual analytics approach to facilitate the proctoring of online exams," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, Yokohama, Japan, 2021.

[18] A. A. Turani, J. H. Alkhateeb, and A. R. A. Alsewari, "Students online exam proctoring: a case study using 360 degree security cameras," in *2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*, pp. 1–5, Bangladesh, 2020.

[19] K. Garg, K. Verma, K. Patidar, and N. Tejra, "Convolutional neural network based virtual exam controller," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 895–899, Madurai, India, 2020.

[20] X. Li, K.-m. Chang, Y. Yuan, and A. Hauptmann, "Massive open online proctor: protecting the credibility of moocs certificates," in *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pp. 1129–1137, Vancouver BC Canada, 2015.

[21] A. H. S. Ganidisastra and Y. Bandung, "An incremental training on deep learning face recognition for m-learning online exam proctoring," in *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, pp. 213–219, Bandung, Indonesia, 2021.

[22] J. Jia and Y. He, "The design, implementation and pilot application of an intelligent online proctoring system for online exams," *Interactive Technology and Smart Education*, vol. 19, no. 1, pp. 112–120, 2022.

[23] M. M. Masud, K. Hayawi, S. S. Mathew, T. Michael, and M. El Barachi, "Smart online exam proctoring assist for cheating detection," in *Advanced Data Mining and Applications: 17th International Conference, ADMA 2021*, pp. 118–132, Sydney, NSW, Australia, 2022.

[24] D. D. Burke and K. J. Sanney, "Applying the fraud triangle to higher education: ethical implications," *Journal of Legal Studies Education*, vol. 35, no. 1, pp. 5–43, 2018.

[25] K. A. Burgason, O. Sefiha, and L. Briggs, "Cheating is in the eye of the beholder: an evolving understanding of academic misconduct," *Innovative Higher Education*, vol. 44, no. 3, pp. 203–218, 2019.

[26] F. B. Withrow, "Technology in education and the next twenty-five years," *THE Journal*, vol. 24, no. 11, pp. 59–61, 1997.

[27] T. Saba, A. Rehman, N. S. M. Jamail, S. L. Marie-Sainte, M. Raza, and M. Sharif, "Categorizing the students' activities for automated exam proctoring using proposed deep l2-graftnet cnn network and aso based feature selection approach," *IEEE Access*, vol. 9, pp. 47639–47656, 2021.

[28] F. Yu, R. Nakashima, K. Matsumiya, I. Kuriki, and S. Shioiri, "Eye-head coordination for visual cognitive processing," *PLoS One*, vol. 10, no. 3, article e0121035, 2015.

[29] E. Kowler, Z. Pizlo, G.-L. Zhu, C. J. Erkelens, R. M. Steinman, and H. Collewijn, "Coordination of head and eyes during the performance of natural (and unnatural) visual tasks," *The Head–Neck Sensory Motor System*, pp. 419–426, 1992.

[30] C. Lee, "Eye and head coordination in reading: roles of head movement and cognitive control," *Vision Research*, vol. 39, no. 22, pp. 3761–3768, 1999.

[31] C. Luo and J. M. Franchak, "Eye-head-body coordination in the motor-memory trade-off," *Collabra: Psychology*, vol. 8, no. 1, article 38821, 2022.

[32] H. P. Graf, E. Cosatto, V. Strom, and F. J. Huang, "Visual prosody: facial movements accompanying speech," in *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, pp. 396–401, Washington, DC, USA, 2002.

[33] E. Z. McClave, "Linguistic functions of head movements in the context of speech," *Journal of Pragmatics*, vol. 32, no. 7, pp. 855–878, 2000.

[34] C. Lugaresi, J. Tang, H. Nash et al., "Mediapipe: a framework for perceiving and processing reality," in *Third Workshop on*

Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR), Long Beach, California, 2019, https://research.google/pubs/mediapipe-a-framework-for-perceiving-and-processing-reality/.

[35] A. Ablavatski, A. Vakunov, I. Grishchenko, K. Raveendran, and M. Zhdanovich, "Real-time pupil tracking from monocular video for digital puppetry," 2020, https://arxiv.org/abs/2006.11341.

[36] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 20, no. 2, pp. 215–232, 1958.

[37] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[38] T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition*, pp. 278–282, Montreal, QC, Canada, 1995.

[39] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.