International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 01 | Jan 2021 www.irjet.net p-ISSN: 2395-0072

AUTOMATING ONLINE PROCTORING THROUGH ARTIFICIAL INTELLIGENCE

Neelesh Chandra M[1], Piyush Sharma[2], Utkarsh Tripathi[3], Ujwal Kumar[4], Dr. G.C. Bhanu Prakash[5]

[1-4]Student, Dept. of Computer Science Engineering, Sir MVIT, Karnataka, India
[5]HOD, Dept. of Computer Science Engineering, Sir MVIT, Karnataka, India

---

Abstract - *In this work, we intend to create a web based automated examination system which can detect any malicious activities and flag them, to ensure fair proposition of exams. Essentially, we detect malpractices by incorporating computer audio and visual movements using webcam and microphone. While it allows students to take a test from any location with specific technical prerequisites, it also removes the need for physical examination centers. Primarily, Vision based tracking consists of Eye ball tracking, Lip movement, Face spoofing, Mobile phone detection, Additional member detection in frame and more. Secondly, the audio-based flagging maps audio to text conversion using Google speech recognition API and also flags if there are high noise disturbances.*

***Key Words*: Proctoring, Vision, Online examination, Malpractice detection, Audio**

1. INTRODUCTION

Remote examination and proctoring are significantly gaining importance in the wake of accommodation of comfort, security and accessibility. This could not just increase importance for course or stream based examinations but also help in MOOCs and other credit-based certifications for the concern of establishing credibility. Instead of taking examinations in a traditional classroom architecture, now we could lay emphasis on comfort-based learning and verification by means of digital proctoring on a remote basis. According to the UNESCO Educational Disruption and Response to COVID-19 pandemic, most of the governments across the globe are closing down educational institutions and are significantly moving their activity to online and remote modality impacting over 89% of the world's student population. Essential emphasis still remains on accuracy of the models used for proctoring and the depth of analysis that takes place to avoid malpractices in case of remote environment which includes high level analogy of suspicious movement detection and flagging and elimination or hardening false positives until significant accuracy is achieved. While this sort of system invites, multiple detection mechanisms, which could be on fronts of Face detection, Noise detection, Eyeball movement detection, change of tabs detection, Device detection and more, often one or more together could facilitate the fairness of examination and add credibility and integrity to it, apart from identity verification to avoid non-repudiation.

2. LITERATURE SURVEY

In [1] published by Asep Hadian S. G and Yoanes Bandung, a unique approach is followed wherein user verification is at high importance on a continuous scale. A large dataset of user images is used to train the CNNs to identify the user in low light and general scenarios. They have achieved this by using filters that detect features that must pass non-linear mapping such that the CNN can learn the values. This system is trained during times of lecture for a MOOC or Classroom setup instead of a first-time approach just during exams, which provides leverage to higher accuracy and larger dataset, which in turn eliminates false positives while training the model. The system accuracy rate is measured in the final evaluation stage using false acceptance rate (FAR) and false rejection (FRR) in the verification process.

In [2] presented by Aiman Kiun, emphasis is laid upon fraud detections in video recordings of exams based on Convolutional Neural Networks (CNN), wherein image classification used models were based on usage of Rectified activation units (RAU) which in turn exhibited amazing results for large scale of data sets. Their framework consisted of three parts, (I) An interface, (II) Video processing, (III) Frame classification. The interface necessarily will send the video of the students taking the examination, into a pipeline that consists of a series of methods. The large recording would be converted to minimalistic frames in number and various duplicate or similar looking frames are eliminated as their candidate frames would already be present in the examining set, this will reduce the burden of the proctoring setup, easing down from hours of content to just few hundreds to thousands of frames. The above exhibits typical optimization for burden reduction and efficient analysis in limited resources. The above frames are then sent into a pipeline where, in the second half of the pipeline, these frames are used to train CNNs to identify the normal and suspicious behavior, based on which further flagging mechanism would come into place in accordance with the various policies and rules that are established to perform the distinguishment. This is the entire process that has been followed in the work presented in this paper.

In [3] by Yousef Atoum, there is an additional proctoring check which has introduced a multimedia analysis system which features gaze detection, phone detection, text

detection, speech detection, covariance feature. They have also used wear cam alongside webcam to facilitate gaze detection. Webcam and gaze cam are both synchronized to provide real time cross check of what the user views. The cheat detection feature here checks various parameters together and in case of entropy across various mechanisms employed such as noise detection, failure of valid gaze detection and other features, it flags the user which is a consequence of eliminating false positives and confirmed cheating cases. A binary SVM classifier is used where the sound frequency is divided into 16 different channels and speech is considered a positive sample and other categories of speech are purely negative samples. Using component analysis of the visible region of the screen, the screen is extracted by computing the convex hull of the large region. Overall cheating behavior detection is carried out by considering output vectors of all the above mechanisms and checking sync of the same to flag activity as malicious.

In [4] by A.T. Awaghade, D. A. Bombe, T. R. Deshmukh and K. D. Takawane, the framework proposed incorporates all the contributions to measure and gauge the assortment of occasions, practices and examples ordinarily connected with cheating. By joining consistent personality check and programmed recognition of misbehavior or dubious exercises by an understudy, this framework gives a versatile, on the web, totally computerized, human communication free delegating framework that can be gotten to by test takers and directors to a really productive answer for customary issue of online test administering. The screen's webcam has an inherent receiver to catch any solid in the room. This framework additionally catches the screen captures from an understudy's machine aimlessly, an ideal opportunity to guarantee respectability. Validation of the personality of the test takers is a significant and possibly costly issue in internet testing. In this framework, confirmation is cultivated utilizing the webcam and straightforward, solid acknowledgment methods.

In [5] by Sanjana Yadav and Archana Singh, they have used computer vision for information extraction for object detection. Computer vision can be used to carry out visual manipulation and recognition as object frames. Steps used by them are as follows, an image is captured and all the objects in the image are stored in a database or a collection. Images are taken at a regular interval which then is matched from the image which is collected in the database. If the position of the object is displaced, then it is flagged separately on similar lines. Image is processed under a matching algorithm using the methods such as rescaling, filtration and binarization. Chamfer distance transformation is used here for detection of the position of the object and map the relativity within the image. This helps us to understand that objects have been moved and similar methods can be used in detection in case of foreign objects in the frame.

In [6] by Swathi Prathish, there are four major channels of detection, (A) System usage analysis, (B) Video analysis, (C) Audio analysis, (D) Inference system. In audio analysis, a very generic system is used where they identify that background noise is always persistent and hence, they train the dataset based on generic noise levels and then calculate the exam dataset and find the ambience based on the calculated average. They have a subsequent rule-based inference system which is driven by rules of flagging which acts upon incoming flags from various channels. This also evaluates multiple faces detection and duration for face disappearance. Yaw angle (the angle by which the face turns left or right) is calculated where up to 43 degrees of movement is acceptable as per rule-based inference, anything more than that is flagged as supposed malpractice.

In [7] by Vahid Kazemi and Josephine Sullivan, the theory tends to the issue of Face Arrangement for a solitary picture. It shows how an outfit of relapse trees can be utilized to gauge the face's milestone positions or landmarks from an inadequate subset of pixel powers, accomplishing execution with excellent forecasts. They present an overall structure dependent on inclination boosting for learning a gathering of relapse trees that advances the number of square losses and normally handles missing or incompletely named information. They show how utilizing fitting priors misusing the design of picture information assists with effective component choice. Distinctive regularization systems and its significance to battle overfitting are likewise researched. Moreover, they examine the dependence of accurate predictions on the quantity of available training data.

In [8] by N.L Clarke and P. Dowland, the paper presents a feasible model to facilitate remote and electronic proctoring during examinations of students. The strategy involves utilizing translucent recognition to give a non- disruptive and persistent authentication of student's identity all through the period of test taking. A model is created and an appraisal of the technology of the created platform showcases the success of this method.

3. CONCLUSIONS

Invigilation or surveilling an exam, is not restricted to a planned time and actual test grounds any longer. The coronavirus pandemic prompted a quick upheaval in the manner the tests are currently being held or will be held in the coming future. This has provoked an interest for an assortment of online distant delegating arrangements and supporting technological advancements. Today, instructors and understudies are progressively receiving virtual platforms to distantly delegate and take the assessments from home bringing about modernizing the plan of leading evaluations. Here, what has not changed is the essential need to adjust security requests with a positive competitor's experience making distant delegating a chance, which is presently generally utilized by both public and private area associations. As explained, the proposed system would be split into two sets of functionalities – vision and audio. The vision-based functionalities would include detecting the eyes, face, mouth and other objects while the voice-based functionalities would record audio from the microphone.

This system can be combined with a secure exam browser to stop cheating from occurring. However, the system will not be completely successful in eradicating all kinds of cheating, so in such cases, human intervention may be required.

REFERENCES

[1] "A Design of Continuous User Verification for Online Exam Proctoring on M-Learning", Hadian S. G. Asep; Yoanes Bandung, 2019 International Conference on Electrical Engineering and Informatics (ICEEI), 9-10 July 2019.

[2] "Fraud detection in video recordings of exams using Convolutional Neural Networks", Aiman Kuin, University of Amsterdam, June 20, 2018.

[3] "Automated Online Exam Proctoring", Yousef Atoum, Liping Chen, Alex X. Liu, Stephen D. H. Hsu, and Xiaoming Liu, IEEE Transactions on Multimedia, 20 January 2017.

[4] "Online Exam Proctoring System", A.T. Awaghade, D. A. Bombe, T. R. Deshmukh, K. D. Takawane, International Journal of Advance Engineering and Research Development (IJAERD) "E.T.C.W", January -2017, e-ISSN: 2348 - 4470, print-ISSN: 2348-6406.

[5] "An Image Matching and Object Recognition System using Webcam Robot", Sanjana Yadav; Archana Singh, 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), 22-24 Dec. 2016.

[6] "An Intelligent System For Online Exam Monitoring", Swathi Prathish, Athi Narayanan S, Kamal Bijlani, 2016 International Conference on Information Science (ICIS), 12-13 Aug. 2016.

[7] "One Millisecond Face Alignment with an Ensemble of Regression Trees", Vahid Kazemi,Josephine Sullivan, 2014 IEEE Conference on Computer Vision and Pattern Recognition,23-28 June 2014.

[8] "e-Invigilator: A biometric-based supervision system for e-Assessments", N.L Clarke, P. Dowland, S.M. Furnell International Conference on Information Society (i-Society 2013), 24-26 June 2013.

[9] "Eye-gaze detection with a single WebCAM based on geometry features extraction", Nguyen Huu Cuong; Huynh Thai Hoang, 2010 11th International Conference on Control Automation Robotics & Vision, 7-10 Dec. 2010.

[10] "Face Tracking and Facial Feature Detection with a Webcam", R. Belaroussi, M. Milgram t, The 3rd European Conference on Visual Media Production (CVMP 2006) - Part of the 2nd Multimedia Conference 2006, 29-30 Nov. 2006.