Detection of Malpractice in Offline Examination Using Deep Learning

[1]Manoj A, [2]Insha Mohammed, [3]Teja Swaroop Naidu, [4]Rohith S R and [5]R Aruna
Department of Electronics and communication Engineering, AMC Engineering college, Bangalore, India.
[1]manoj181297@gmail.com, [2]inzy2303@gmail.com, [3]ntejaswaroop@gmail.com, [4]aruna.ramalingam@gmail.com

Article Info
Jenitta J and Swetha Rani L (eds.), *International Conference on VLSI, Communications and Computer Communication*, Advances in Intelligent Systems and Technologies,
Doi: https://doi.org/10.53759/aist/978-9914-9946-1-2_29
©2023 The Authors. Published by AnaPub Publications.

Abstract - Exam proctoring is a hectic task i.e.; the monitoring of students' activities becomes difficult for supervisors in the examination rooms. It is a costly approach that requires much labor and difficult task for supervisors to keep an eye on all students at a time. Automatic exam activities recognition is therefore necessitating and a demanding field of research. In this research work, categorization of students' activities during the exam is performed using a deep learning approach. Adeep CNN architecture a kernel size of 7 * 7 and 64 different kernels all with a stride of size 2 givingus 1 layer. After that, the model is validated upon ImageNet. In this paper, we present amultimedia analytics system which performs automatic offline exam proctoring. The system hardwareincludes one webcam for the purpose of monitoring the visual environment of the testing location. Toevaluate our proposed system, we collect multimedia (visual) data from many exam centers performing various types of activities while taking exams. Extensive experimental results demonstratethe accuracy, robustness, and efficiency of our offline exam proctoring system.

Keyword: CNN, ResNet50, SVM

I. INTRODUCTION

Students action modelling a new field of study. It works on learning `how' and `why' of activities that take the human view of knowledge. [1] In any instructional method, evaluation plays a key part as a means of measuring the students' understanding of the topics associated with educational outcomes. Student deception through real-time examinations, regardless of the degree of advancement of the community, is a common issue across the globe. [2] There seems to be no widely held description of deception in exams, as with the principle of unprofessional conduct, but many types of actions are normally considered as necessary indicators of student academic deception, all of which have in particular that they are unethical acts aimed at enhancing the perceived success.

The conventional exam inspection in institutions is based on a written paper and involves the invigilators'appearance on the spot. Figure 1 shows two different scenes of exam activities. [3] The invigilators wander around desks to keep checking for any ethical activity by students. Manual tracking is a challenging method for assessingthe behavior of students during written exams. Because of intellectual dishonesty, many students have capturedannually.

Fig 1. Examination Hall

In tests, [4] deception means violating the exam rules. In a multitude of ways, students trick the examiner, including covering notes, transmitting hidden messages by hand and body movements, monitoring nearby student documents, passing notes to other students while the examiner is not watching, and illegally using technological devices such as mobile phones, calculators and digital watches. There are several explanations for the deceit of students in examinations. These causes are greatly divided into (a) psychiatric and (b) social problems**.**

Psychiatric [5] difficulties include a lack of time and anxiety factors in examinations. Family peer pressure and feeling incompetent are linked to psychiatric disorders. Exam deception is implicated in the pathogenesis of potential dishonest conduct, both in higher education and subsequently in professional life. Also, the use of unethical activities in exams can lead to serious implications for both individuals and society.

Some of such implications contain: [6] (a) In one social background, developing a deceptive habit is therefore prone to leaking over to one another, (b) The idea that stealing contributes to an erroneous competence appraisal of the student often implies that the person's basic requirements for learning skills are negatively impacted, and (c) reduction in reputation of the corresponding organization.

To preserve [7] the student's integrity actively during the assessment process, real-time surveillance is very needful. This can help in the reduction of invigilators' tough duty. Also, it will be difficult for the students to dodge the camera in front of the students as they can do when the invigilator sees towards other direction.

A test or examination (exam) is an assessment [8] intended to measure a person’s knowledge, skill, or ability in one specific subject or a variety of topics and in present day are primarily administered on paper or on a computer. This can be done in-person or online. A proctor is a person who takes charge of, or acts for, another. The word proctor is frequently used to describe someone who oversees an exam. In today’s testing environment, a proctor will typically verify a student’s activities and identity. Exam proctoring application is a need of the hour that ensures health safety while also eliminating the chances of cheating or malpractice during examinations. To promote fairness in examinations. To deter corrupt practices in examinations.

II. RELATED WORK

They [9] did the evaluation to verify the results in three main phases: student ID, the position of the student (via row and column), and gaze. First, it is the student ID that needs to be detected primarily. The student IDs maintain an important role in this context. Once all student IDs have been identified and located, the tracked dataof individuals’ behaviors will be attributed to them later. Student ID identification is evaluated through all the data of the dataset. Because the data are imbalanced, F1-scores are necessary. [10] A confusion matrix is also plotted.The first column and row represent the label of “unknown,” and the other columns and rows show the results of corresponding student IDs. Secondly, row and column are evaluated. The row and column represent the current position of the student in the class which is going to be combined with the head-pose direction to denote the originand the direction of the gaze vector. [10] The row and column are evaluated with MAE (mean absolute error). Besides,confusion matrices are also constructed for those estimations; vertical and horizontal values of the matrices are matched with the ranges of parameters. Finally, the gaze plays the most pivotal role in the system, to check if thestudents are focusing on the board/slides, on laptops, or on other things. The summarized statistics of gaze couldbe exhibited for educators to observe the behaviors of attention over the studying period. Gaze estimation is acquired through re-trained models. Hence, the dataset is divided into training and testing sets, and then one-thirdof the dataset (7556 rows) is used for evaluation. The F1-score is also applied to evaluate the result of gaze estimation.

This [11] paper presents a multimedia analytics system for online exam proctoring, which aims to maintain academic integrity in e-learning. The system is affordable and convenient to use from the text taker’s perspective,since it only requires to have two inexpensive cameras and a microphone. With the captured videos and audio, we extract low-level features from six basic components: user verification, text detection, speech detection, activewindow detection, gaze estimation and phone detection. These features are then processed in a temporal windowto acquire high-level features, and then are used for cheat detection. Finally, with the collected database of 24 test takers representing real-world behaviors in online exam, we demonstrate the capabilities of the system, with nearly 87% segment-based detection rate across all types of cheating behaviors at a fixed FAR of 2%. These promising results warrant further research on this important behavior recognition problem and its educational application.

This [12] work is based on a proposed CNN network i.e. L2-GraftNet for feature extraction along with the ASO algorithm for feature selection. The proposed L2-GraftNet is first converted to a pre-trained model by performing training on the CIFAR-100 dataset. The features are taken from the pre trained model on the CUI- EXAM dataset. The CUI-EXAM dataset is first annotated and augmented before taking its image features. Thesefeatures are forwarded to ASO based features optimization method and through various classifier variants of SVMand KNN, the model performance is observed on the selected features. 5-Fold cross-validation is applied for testing and training of the dataset. Many experiments are performed, apart, only five experiments are discussed in detail. It is observed that the least performance outcomes are obtained for experiments with 100 features havingan accuracy of 92.43 percent by using the FKNN classifier. [13] Similarly, the best performance outcomes are considered for experiments with 1000 features having an accuracy of 93.88 percent by using the FKNN classifier.From all classifiers' outcomes, in all experiments, it is observed that FKNN shows high performance and CSVMdepicts second to high performance in terms of selected performance measures. Although the proposed system shows satisfactory outcomes in terms of performance measures, still there can be improvements performed for further enhancement of accuracy. In the future, the work can further be explored with more state-of-the-insight approaches such as manifold learning, LSTMs, Quantum deep learning, and brain-like computing Approaches.

Online (4) teaching still does not offer complete remote teaching in most cases, since there are many institutions that, in the evaluation process, continue to require the physical presence of the student in a specific place to unite the student and the examiner in said place, for supervisory reasons. [14] However, there are already e- proctoring tools that allow this process to be carried out remotely, without requiring that physical presence. Additionally, there is a favorable trend in the application of this methodology in MOOCs and in open education globally. Thus, this study has sought to locate the motivational factors determining the implementation of this evaluation system, allowing the exposition of a list of motivational influencing factors when accepting the use of new technological tools (that is, when the educational system accepts this tool as a method of remote supervision),and determine which are the most influential or decisive when it comes to acceptance by educational institutions. The list is made up of the following motivational factors: [15, 16] Quality management (QM), available information (AI), external conditioning (EC), trust (T), perceived compatibility (PC), perceived usefulness (PU), attitude (A) and intention (I). The most decisive factor in this process is trust (T), which would be the degree of security and privacy that institutions have in the use of this tool (e-proctoring). This coincides with the main line of research on this tool, where most of the studies focus on trust and the safety of using it. Something that can be seen in studies from years ago, as can be seen in Howlett and Hewett, is that both technological solutions and instructionaldesign solutions to reduce cheating in these remote examinations were already examined

Auto proctoring process (Fig 2) can automate exam invigilation and candidate's activities while conducting offline exams. The invigilators can sit at a location which is far away from the candidate and control the exam. They can get the alert based on their activities as also supervise them remotely. Offline exams can be easily conducted without worrying about cheating and malpractices. It enables the candidates to appear for the proctoring offline exam from any location. Thus, Auto proctoring plays a significant role in eliminating the location and manpower constraints associated with the exam center.

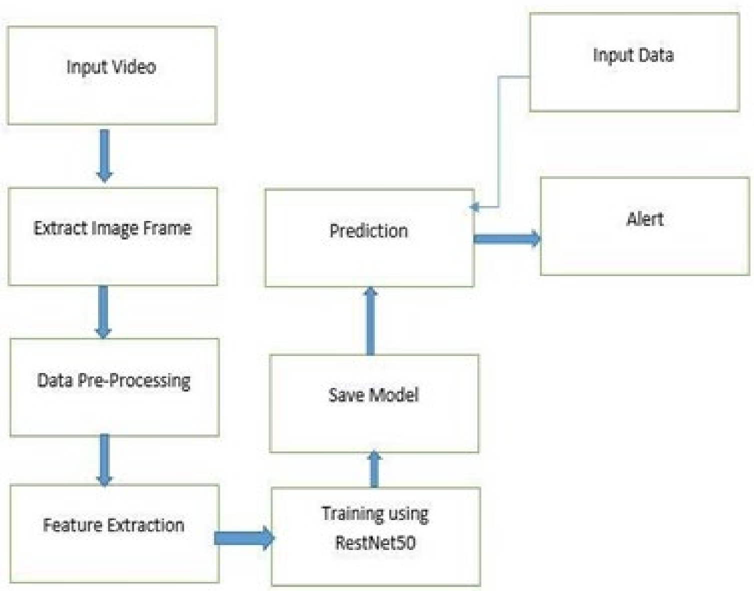

Fig 2. Block Diagram

The main objectives are

- To develop a trained model from the real time data of the exam hall.
- To train the deep learning model based on different activities during exam.
- To monitor the activities of the students during offline exam programmatically using AI.

The architectural configuration procedure (Fig 3) is concerned with building up a fundamental basic system for aframework. It includes recognizing the real parts of the framework and interchanges between these segments. Thebeginning configuration procedure of recognizing these subsystems and building up a structure for subsystem control and correspondence is called construction modeling outline and the yield of this outline procedure is a portrayal of the product structural planning. The proposed architecture for this system is given below. It shows the way this system is designed and brief working of the system.

Dataflow: Movement of data is shown by pointed arrows. Data movement is shown from the base of arrow asits source towards head of the arrow as destination.

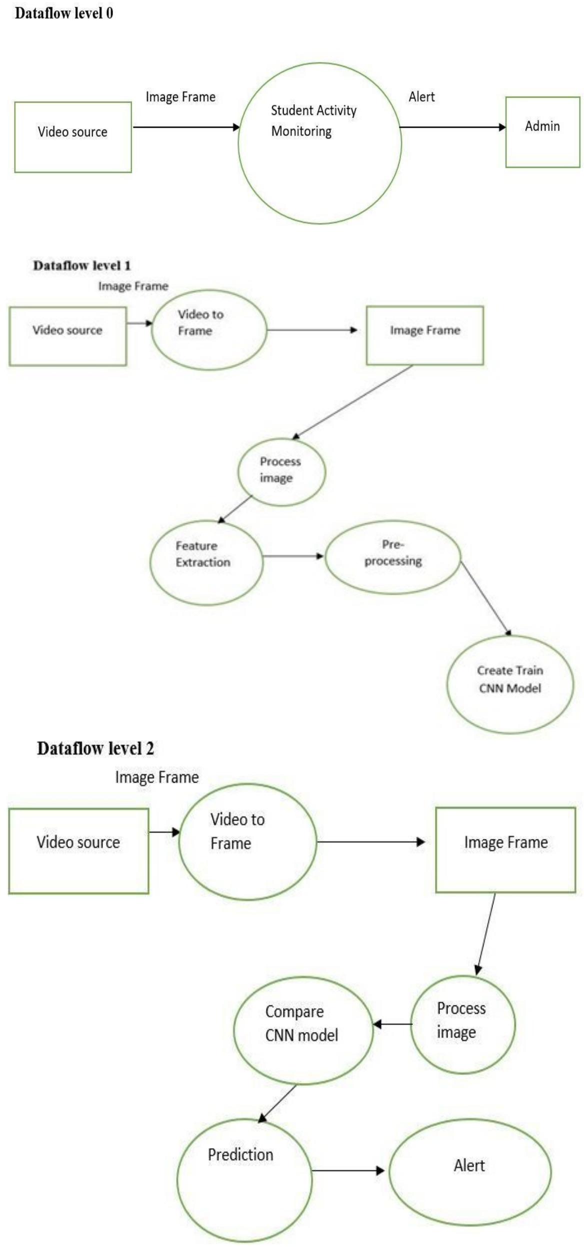

Fig 3. Dataflow level 0,1,2

Split video: As the size of the video is huge and in one video multiple activities are involved, the video has been divided into multiple pieces using moviepy python module. We have mentioned the start and end time of each video chunk. The chunk videos are named sequentially.

Extract images: To extract the images from the selected video, we have used OpenCV library and to crop the images we have used PIL. The image has been cropped to remove the unwanted part of each image.

Initial training and generation: In our proposed work, the multi-scale image block-level CNN object detection model is proposed. In the training phase, the input image of this method should satisfy two conditions. One condition is that the image requires image block-level label. The other condition is that the image contains different sizes. So, all the images are converted into same size images. The training data is divided into two parts, positive samples (objects) and negative samples (backgrounds). Positive training samples are the square image block which contains the single object.

Data Augmentation: We all know that deep learning model requires a large amount of data to train. Therefore, thepositive sample should be augmented. First, the center offset and the range expansion is processed foreach positive sample. The center offset includes top-left, top-mid, top-right, left-mid, right-mid, bottom-left, and bottom-mid, bottom-right. The ranges of the center offset and the range expansion are one-tenth of the original positive sample edge length. Second, the original data and augmented data are expanded by rotating 90◦, 180◦, 270◦. Third, the original data and augmented data generatedby several forward steps are expanded by mirroring.

Multi-scale image block-level CNN: In this subsection, the multi-scale image block-level fully convolutional neural network will be described in detail. The deep learning model has become a very important tool in the field of imageprocessing. The convolutional neural network is one of the most representative models. In this work,the multi-scale and image block-level fully convolutional neural network consists of multiple branches each of which contains convolution layer, rectified linear unit (ReLU) layer, local responsenormalization layer, and SoftMax layer.

Activity detection: In this subsection, the detection process of the proposed method will be described in detail. It contains feature extraction and post processing. In the testing phase, these parts are different from the training phase.

Algorithm:

CNN Algorithm overview: Convolutional Neural Network (CNN) were used to achieve some breakthrough results and win well-known contests. The application of convolutional layers consists in convolving a signal or an image with kernels to obtain feature maps. So, a unit in a feature map is connected to the previous layer through the weights of the kernels. The weights of the kernels are adapted during the training phase by back propagation, in order to enhance certaincharacteristics of the input. Since the kernels are shared among all units of the same feature maps, convolutional layers have fewer weights to train than dense FC layers, making CNN easier to train and less prone to overfitting. Moreover, sincethe same kernel is convolved over all the image, the same feature is detected independently of the locating— translation invariance. By using kernels, information of the neighborhood is taken into account, which is a usefulsource of context information. Usually, a non-linear activation function is applied on the output of each neural unit. If we stack several convolutional layers, the extracted features become more abstract with the increasing depth.

Architecture:

Resnet 50: ResNet50 (Fig 4) is a variant Resnet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. The ResNets were initially applied to the image recognition task but as it is mentioned in the paper thatthe framework can also be used for non-computer vision tasks also to achieve better accuracy.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Fig 4. Resnet Architecture

Support Vector Machines: A Support Vector Machine (SVM) (Fig 5) is a supervised machine learning algorithm that can be employed for both classification and regression purposes. SVMs are more commonly used in classification problems and as such, this is what we will focus on in this post. SVMs are based on the idea of finding a hyperplane that best divides adataset into two classes, as shown in the image below.

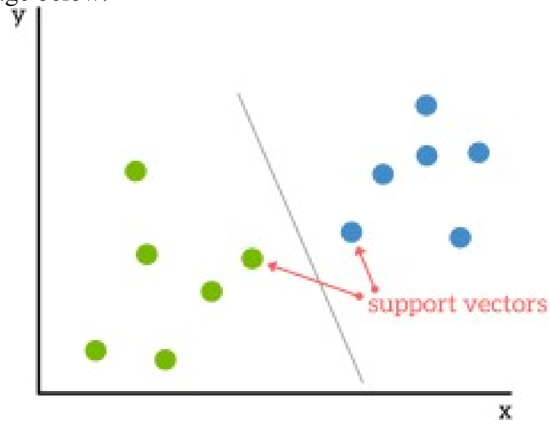

Fig 5. Hyperplane in SVM

III. EXPERIMENTAL RESULT

Fig 6. Detected Peeking

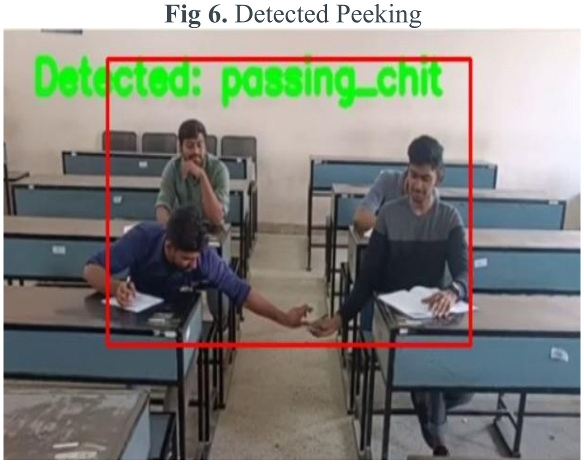

Fig 7. Detected Passing

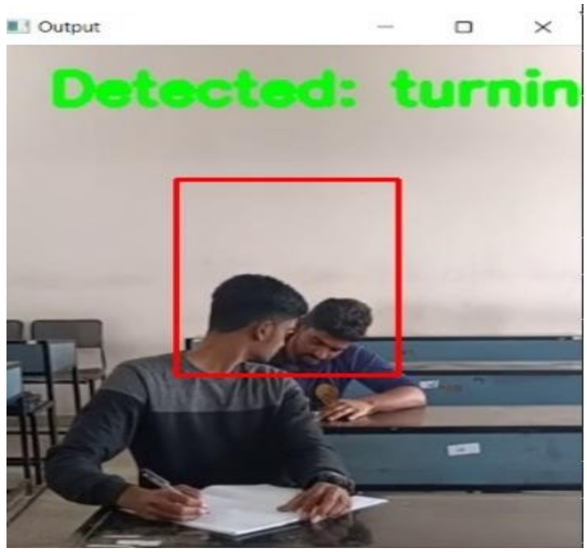

Fig 8. Detected Turning Back

Our proposed system validates the user authentication and detects the misbehaviour of the examinee (Fig 6, Fig 7, Fig 8). Those are done by using deep learning with high accuracy. We hope to do something more than the research area we are currently doing and integrate this one into a very good web-based software that combines all the features that a single online system should have. And to go beyond the ongoing studies and create a system withthe highest level of accuracy using AI technology (we hope to use AI technology for improving voice recognition of users). At the same time, our next step is to connect the mobile phone and connect the user to thissystem. We also hope to create a related mobile app.

V. CONCLUSION

This paper designs a set of solutions for misbehavior monitoring of offline examination based on video processing and work is based on a proposed CNN network i.e., ResNet50 for feature extraction along with the feature selection techniques. The proposed ResNet50 is first converted to a trained model by performing trainingon the real time offline classroom dataset. The features are taken from the trained model on the real time offline classroom dataset from our college. The dataset is first annotated and augmented before taking its image features.5-Fold cross-validation is applied for testing and training of the dataset. This method will help us to monitor the activities of the students during offline exam Which gives a high accuracy rate and the best performance. Data gathering and evaluating the accuracy was based on previous research. In this offline examination system, the user will be monitored by the system throughout the exam. This is core and this is a very critical part. The fully completed system would be able to be used for higher educational institutions and many more.

References

[1] A Lam, Phan & Phuong Chi, Le & Tuan, Nguyen & Dat, Nguyen & Nguyen, Trung & Anh, Bui & Aftab, Muhammad Umar & Tran, Van Dinh & Son, Ngo. (2019). "A Computer-Vision Based Application for Student Behavior Monitoring in Classroom". Applied Sciences. 9. 4729. 10.3390/app9224729.

[2] B. Atoum, Yousef & Chen, Liping & Liu, Alex & Hsu, Stephen & Liu, Xiaoming. (2017). "Automated Online Exam Proctoring". IEEE Transactions on Multimedia. PP. 1-1. 10.1109/TMM.2017.2656064

[3] 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," in IEEE Access, vol. 9, pp. 47639-47656, 2021, doi: 10.1109/ACCESS.2021.3068223.

[4] González González, C. S., Infante Moro, A., & Infante Moro, J. C. (2020). Implementation of E-Proctoring in Online Teaching: A Study about Motivational Factors. Sustainability, 12(8), 3488. DOI: https://doi.org/10.3390/su12083488.

[5] A. R. Baig and H. Jabeen, ``Big data analytics for behavior monitoring of students," *Procedia Comput. Sci.*, vol. 82, pp. 43_48, Jan. 2016.

[6] F. Rodrigues and P. Oliveira, ``A system for formative assessment and monitoring of students' progress,"

[7] *Comput. Educ.*, vol. 76, pp. 30_41, Jul. 2014.

[8] J. Ramberg and B. Modin, ``School effectiveness and student cheating: Do students' grades and moral standards matter for this relationship?" *Social Psychol. Educ.*, vol. 22, no. 3, pp. 517_538, Jul. 2019.

[9] Z. A. von Jena, ``The cognitive conditions associated with academic dishonesty in university students and itseffect on society," *UC Merced Undergraduate Res. J.*, vol. 12, no. 1, pp. 1_21, 2020.

[10] M. Ghizlane, B. Hicham, and F. H. Reda, ``A new model of automatic and continuous online exam monitoring," in *Proc. Int. Conf. Syst. Collaboration Big Data, Internet Things Secur. (SysCoBIoTS)*, Dec. 2019, pp. 1_5.

[11] I. Blau and Y. Eshet-Alkalai, "The ethical dissonance in digital and nondigital learning environments: Does technology promotes cheating among middle school students" *Comput. Hum. Behav.*, vol. 73, pp. 629_637, Aug.2017.

[12] A. Asrifan, A. Ghofur, and N. Azizah, ``Cheating behavior in EFL classroom (a case study at elementary school in Sidenreng Rappang Regency)," *OKARA, J. Bahasa dan Sastra*, vol. 14, no. 2, pp. 279_297, 2020.

[13] P. M. Newton, ``How common is commercial contract cheating in higher education and is it increasing? A systematic review," *Frontiers Educ.*, vol. 3, p. 67, Aug. 2018.

[14] A. Bushway andW. R. Nash, ``School cheating behavior," *Rev. Educ. Res.*, vol. 47, no. 4, pp. 623_632, Dec.1977.

[15] Brown, M. T. (2017). Automated Grading of Handwritten Numerical Answers. In 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR) (pp. 279-284).

[16] IEEE Cupic, M., Brkic, K., Hrkac, T., Mihajlovic, Z., & Kalafatic, Z. (2014, May). Automatic recognition of handwritten corrections for multiple choice exam answer sheets. In Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on (pp. 1136-1141).