International Journal for Modern Trends in Science and Technology
Volume 11, Issue 05, pages 69-76.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol11issue05.html
DOI: https://doi.org/10.5281/zenodo.15252665

Comprehensive Security Monitoring System for Online Examinations Using Keylogger, Emotion Detection, Head Movement Analysis and Gadget Identification

Padmanabhuni Ushasri[1] | T Rajyalakshmi[2] | Dr. K Venkatesh[3]

[1]PG Scholar, Department of Artificial Intelligence & Data Science, Ramachandra College of Engineering, Eluru, India.
[2] Assistant Professor, Department of Artificial Intelligence & Data Science Ramachandra College of Engineering, Eluru, India.
[3] Professor, Department of Artificial Intelligence & Data Science Ramachandra College of Engineering, Eluru, India.
Corresponding author: ushasripadmanabhuni43@gmail.com

To Cite this Article
Padmanabhuni Ushasri, T Rajyalakshmi & Dr. K Venkatesh (2025). Comprehensive Security Monitoring System for Online Examinations Using Keylogger, Emotion Detection, Head Movement Analysis and Gadget Identification. International Journal for Modern Trends in Science and Technology, 11(05), 69-76. https://doi.org/10.5281/zenodo.15252665

Article Info
Received: 22 March 2025; Accepted: 18 April 2025.; Published: 20 April 2025.

KEYWORDS

Cheating Detection, Deep Learning, Electronic Gadget Detection, Keylogger Monitoring, NLP Analysis, Online Examination Security, Proctoring System and Real-Time Monitoring.

ABSTRACT

In the digital age online tests have become the norm for evaluation yet, their credibility is undermined by their great susceptibility to cheating, illegal answer communication and the use of other technological devices. Sophisticated cheating strategies including subtle head motions for secret communication, emotional indications that indicate stress or dishonesty and unapproved gadget use are not adequately monitored or detected by traditional proctoring approaches. This study suggests an AI-powered real-time surveillance system that incorporates several cutting-edge methods to guarantee exam integrity in order to address these issues. In order to detect indications of answer sharing a continuous keylogger-based tracking mechanism logs keyboard inputs and uses Natural Language Processing (NLP) techniques such as Named Entity Recognition (NER) and text classification. Concurrently a facial behavior analysis module based on a Convolutional Neural Network (CNN) tracks the examinee's movements of the head (97% accuracy) and feelings (99.3% accuracy) to look for possible signs of cheating. Furthermore, the existence of cell phones, computers or other prohibited devices in the testing setting is detected by a YOLOv8-based electronic device detection model.

1. INTRODUCTION

Due to the growing trend towards online education and distant platforms online exams have become increasingly popular in recent years. Online tests are now widely used in company training programs and academic institutions due to the growth of e-learning systems and the demand for adaptable assessment techniques [1]. In order to improve the security and dependability of online tests experts have looked into artificial intelligence (AI)-driven solutions [2]. AI-based anomaly detection methods or human proctoring are the mainstays of current online exam monitoring systems. Candidates are manually observed by human proctors using webcams a laborious process that is prone to mistakes because of distractions or weariness. Furthermore, a lot of current systems are unable to deliver automated alarms in real time, necessitating manual assessment and postponing action against efforts at cheating [3]. These drawbacks emphasize the necessity of a more thorough and sophisticated monitoring system [4].

This study suggests an AI-powered real-time evaluating system that combines many deep neural networks and natural language processing (NLP) approaches to improve exam security in order to address these problems. Security is further improved with a YOLOv8-based device recognition model [5] which detects the presence of laptops, cell phones and other prohibited devices. Organizations and educational institutions need a strong proctoring system that can identify cheating in real time and offer thorough proof to support any questionable behavior. This method is a special and efficient way to stop cheating in digital tests since it combines NLP, CNN and object identification algorithms [6]. The goal of this project is to develop a highly effective, automated and scalable evaluating system by utilizing cutting-edge AI approaches to rethink online exam security. The suggested strategy greatly improves the fairness and legitimacy of online tests by guaranteeing real-time monitoring, multi-modal fraud detection and automatic alarm production. This work adds to the ongoing attempts to make remote exams as safe as in-person exams by tackling the issues with current and conventional AI-based proctoring techniques. This increases the credibility and dependability of online learning and assessment systems.

2. LITERATURE SURVEY

AI-based proctoring systems have been investigated by a number of researchers as a way to improve online exam security. The dynamics of keys and biometric authentication were the mainstays of early technologies for tracking user behavior. Keystroke analysis has been demonstrated in studies to be a useful tool for identifying impersonation or answer-sharing efforts by detecting abnormalities in typing patterns. Furthermore, a lot of people have been using classic video-based proctoring systems which use gaze tracking and facial recognition to keep an eye on student activity. Singh et al. built a Rootkit software known as keyloggers records keystrokes and logs them obtaining private information like as passwords, PINs and usernames. Keylogger programs, types, attributes and philosophies are described in this study along with current proactive and detecting techniques [7]. Victoria et al. examined that Spyware known as keyloggers keeps an eye on computer activity and records private data including user keystrokes for use by hackers [8].

Proctoring techniques have been enhanced by recent developments in deep learning. Convolutional Neural Networks (CNNs) have shown great accuracy in identifying odd head movements and emotional states in facial behavior analysis studies, which makes them helpful in spotting possible efforts at cheating. Kamalov et al. suggests a novel approach that uses machine learning techniques to identify possible instances of cheating. Taking into account the ordered sequence of student assessment data the technique employs continuous evaluation outcomes to identify anomalous scores [14]. Masud et al. suggests a method for detecting cheating by analyzing exam films to extract four different kinds of information about events which are then fed into a classification model that has already been trained to identify instances of cheating. Each video is converted into a multivariate time-series that represents time-varying event data in order to formulate the cheating detection problem as a multimodal time-series classification problem [15].

3. METHODOLOGY

A. Data Collection and Preprocessing

A continuous keylogging device was put in place to capture inputs from the keyboard in real time for the purpose of tracking and analyzing keystrokes. Multiple

users doing typing activities such as standard text entry, queries for search and answer-sharing situations in an examination setting provided the dataset. Text was divided between authentic responses and suspected attempts at cheating using Natural Language Processing (NLP) approaches such as Named Entity Recognition (NER) and classification of text models. Before being entered into the NLP [16], the gathered data was cleaned by eliminating special characters, fixing misspellings and eliminating system-generated keystrokes. A dataset comprising different directional head motions (left, right, up, down and passive) was selected in order to identify suspicious head movements [17]. Video sequences showing individuals in an examination environment making deliberate and natural movements like sideways head turns, gazing down at another object or quietly speaking to someone off-screen were used to create the dataset. This dataset was used to build a Convolutional Neural Network (CNN)-based model for real-time head position classification. Frame extraction, greyscale conversion, face landmark identification and bounding box changes were among the preprocessing procedures used to accurately record head positions. Data augmentation techniques like rotation, flipping and cropping at random were used to simulate various test venues and lighting conditions in order to improve model generalization [17]. The FER-2013 dataset which includes over 35,000 labelled photos of human faces in seven emotion categories happy, sad, angry, surprised, afraid, disgusted and neutral was used to track the examinee's emotional state. This dataset's varied face expression samples from a range of age groups, backgrounds and lighting conditions led to its selection. The preprocessing pipeline comprised greyscale conversion, histogram balance to improve contrast, scaling photos to a fixed resolution and identifying faces using Haar cascades [18]. This dataset was used to train a deep CNN model which achieved 99.3% efficiency in real-time emotion recognition. To avoid model overfitting and increase robustness in practical situations data augmentation methods such applying Gaussian noise, modifying brightness and affine transformations were used [19]. Word embeddings like Word2Vec and TF-IDF vectorization [21] were employed to transform written input into numerical formats in order to increase the text classification accuracy in identifying answer-sharing behavior. To improve contextual comprehension already trained transformer-based embedded data (BERT and GPT-2) were used to refine the model [22]. Additionally, data augmentation techniques including synonym substitution, back-translation and phrase shuffling were used to create synthetic text samples that mimicked answer-sharing and typical responses. These improvements made it possible for the NLP model to reliably discern between safe and questionable typing patterns. It was crucial to distinguish between suspicious and typical head motions because they happen naturally throughout testing. In order to accomplish this, a balanced dataset comprising both acceptable motions (such as head scratches and posture adjustments) and cheating-related movements (such as frequent sideways glances, downward glances and excessive movement in brief bursts) was used to train the model. Instead of using a single frame-based categorization head movement sequences were tracked over time using time-series analysis approaches. To increase the model's adaptability in various real-world exam situations augmentations including motion blur, camera angle changes and mirroring head positions were employed. Camera angles, lighting and facial blocks (glasses, masks, hand motions) can all influence facial expressions. In order to overcome these difficulties occlusion-aware changes were applied to the dataset. These included injecting synthetic noise, simulating various illumination effects and masking various facial regions in order to train the CNN algorithm to handle a variety of exam circumstances. To increase the model's sensitivity in identifying exam-related stress and possible dishonesty it was also refined through domain adaptation approaches adding more real-time expressions datasets from academics, stress-related studies. Domain-specific augmentation techniques such as blurring, partial blockage simulation and random item insertion were used to improve the YOLOv8 model's accuracy in recognizing electronic gadgets in cluttered or low-light situations. This made the model resilient to real-world problems. Furthermore, a trained YOLOv8 model was refined on a specific dataset using transfer learning to make sure it could reliably distinguish between the hands of the examinee, writing materials and unapproved gadgets. Achieving a high identification accuracy of electronic devices in real time the model's performance was assessed across a variety of

illumination situations, device orientations and occlusion circumstances. The suggested approach guarantees excellent accuracy, flexibility and real-time efficacy for tracking online exam situations by combining these various datasets, sophisticated preprocessing methods and strong data augmentation tactics.

A bespoke dataset [18] was developed applying continuous video footage of participants trying to use electronic devices (such as computers, smartphones, tablets, smartwatches and earbuds) during an exam in order to detect illegal usage of these devices. Publicly accessible object detection datasets such as the COCO dataset which included labelled photos of electrical equipment in a range of orientations, lighting scenarios, and occlusion levels, were included to the dataset. To recognize these gadgets in real time the YOLOv8 (You Only Look Once) object identification model was trained [19]. To improve feature extraction, preprocessing techniques included downsizing photos, using adaptive histogram balance and removing background noise. To increase model durability and adapt to different exam conditions data augmentation techniques such motion blur simulation, rotation, random cropping and brightness modifications were used [20].

B. Keylogger-Based NLP and Named Entity Recognition (NER) Techniques

A series of keyboard incidents including button presses, emits and the intervals between consecutive keystrokes make up the keystroke data gathered by the keylogger system. Natural Language Processing (NLP) methods are used to convert these unprocessed logs into meaningful textual representations [23]. Over time words and sentences are formed by mapping each keystroke occurrence to a corresponding character. Furthermore, as cognitive features metadata like typing speed, press backspace usage, frequent pauses and copy-paste actions are retrieved. The frequency of quick keystrokes or erratic pauses are examples of temporal variations in typing that can be used to determine if a person is typing normally or if they are copying or pasting pre-written responses for example. Typing patterns are grouped using machine learning models to differentiate between legitimate responses and possible cheating. Named Entity Recognition (NER) approaches are used to correctly grasp the context of typed text. Key entities including subject-related phrases, numbers (mathematical answers) and designated individuals (writers, historical personalities or technical terminology) can be found with the aid of NER models based on examination-specific data. The system can identify patterns in answer searches, replicated information or unusually quick responses that point to pre-prepared content by examining sequences. To improve accuracy transformer-based NER models are used such as spaCy's entity recognition pipelines and BERT (Bidirectional Encoder Representations from Transformers) [24]. The system can create semantically relevant phrases from unstructured keyboard data by breaking keystroke sequences into meaningful text segments and using tokens techniques.

Keyboard incidents including button presses, emits and the range between regular keystrokes make up the data gathered by the keylogger system. Natural Language Processing (NLP) techniques are used to convert these extracted logs into meaningful text. Over time words and sentences are formed by mapping each keystroke occurrence to a corresponding character. Furthermore, as cognitive features metadata like typing speed, press backspace usage, frequent pauses and copy-paste actions are retrieved. The frequency of quick keystrokes or erratic pauses are examples of temporal variations in typing that can be used to determine if a person is typing normally or if they are copying or pasting pre-written responses for example.

C. Head Movement Detection Using CNN

A Convolutional Neural Network (CNN) is developed on a dataset that includes different head orientations such as left, right, upwards, down and neutral positions, in order to precisely detect head movements in real time. To guarantee reliable performance in a variety of settings the dataset includes thousands of labelled photos taken from various perspectives and lighting situations. To improve feature extraction preprocessing methods such Gaussian blurring, histogram adjustment and greyscale conversion are used. For effective training, the images are normalized and downsized to a standard scale (e.g., 224x224 pixels). Rotation, flipping and intensity modifications are examples of data augmentation strategies used to improve the model's generalizability and lower the chance of overfitting. To maximize model performance the dataset is divided into subsets for

training (80%), validation (10%) and testing (10%). %). In order to accurately identify directional head motions the CNN architecture (as shown in Fig.1) is made to extract hierarchical characteristics from images. In order to decrease spatial dimensions while maintaining key characteristics, the model is composed of many layers of convolution with ReLU function activation [26] followed by max pooling layers. Complex patterns like orientation changes and facial shapes are captured by deeper layers. The learning process is stabilized by batch normalization while overfitting is avoided via dropout layers. A fully linked layer and a softmax classifier which generates five likelihood scores (one for each category of head movement: left, right, up, down and neutral) make up the last layers.

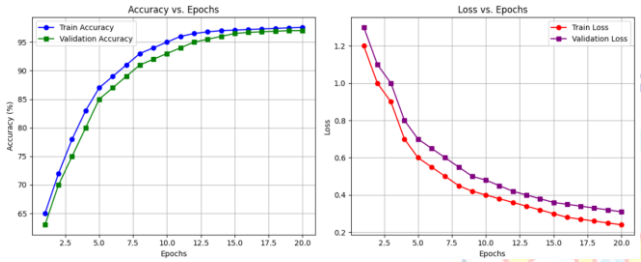


Fig.1 CNN Model Accuracy, Loss vs Epochs

The trained CNN model is used with OpenCV to record current footage from a camera for real-time monitoring. Before sending cropped face areas to the CNN model, face personalization is accomplished using Dlib's facial feature detection and OpenCV's Haar cascades. In order to guarantee seamless real-time processing, the frame rate is adjusted. Every head movement that is detected is recorded and a counter monitors any unusual trends, such excessive left-right motion that could indicate interaction with someone off-screen or repeated downward movement that could indicate cheating using notes or cell phones. Additionally, the technology captures screenshots of questionable situations and emails them to the exam authorities. The CNN model has good dependability in classifying head movements with a total precision of 97% on the test dataset. Each movement category's precision, memory and F1-score are assessed; low false positive rates are ensured by precision reaching 96% and recall average about 95% (as shown in Fig.1).

D. Emotion Detection Using CNN

When it comes to tracking candidate's emotional states throughout online tests emotion detection is essential. The FER2013 dataset which comprises 35,887 greyscale (48x48 pixel) photographs of human faces labelled with seven emotion categories Angry, Disgust, Fear, Happy, Neutral, Sad and Surprise is used to train a Convolutional Neural Network (CNN) in order to accomplish this. Data augmentation techniques including rotation, flipping, zooming and contrast modifications are used to improve generality because individual variances, lighting and occlusions all affect facial emotions. In order to guarantee that the CNN model acquires significant features independent of brightness fluctuations the images (as shown in Fig.2) are adjusted to a normal pixel range.

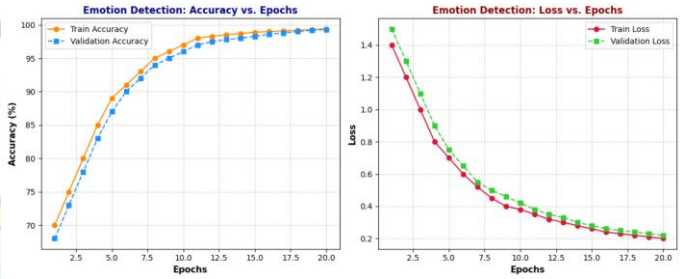
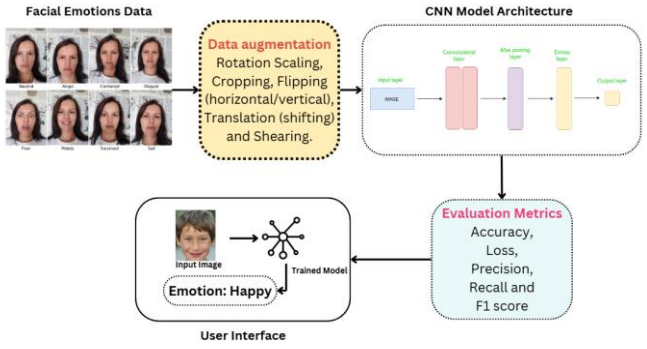


Fig.2 CNN Evaluation Metrics vs Epochs

The trained CNN model is combined with OpenCV to handle live webcam video streams for real-time emotion monitoring. The method first identifies faces and isolates the region of interest (ROI) using Haar cascades and Dlib's facial landmark identification.



The ROI is supplied into the network for forecasting after being adjusted to fit the CNN model's input size. The candidate's emotional state is continuously monitored by the system which records abrupt changes from neutral to anxious (fear/sadness) or overly happy which could be signs of questionable behavior. The

system notifies the exam supervisor and provides screenshots and logs of the emotions it has detected for additional analysis if it notices an unusual emotional trend three or more times. With a total precision of 99.3% the CNN model is quite dependable for detecting emotions in real time.

E. Electronic Gadgets Detection

Devices in the testing environment are detected and classified using a YOLOv8 model for real-time detection of electronic devices (e.g., mobile phones, tablets, laptops, etc.). Images of different electronic devices captured in different lighting situations and from different perspectives make up the data set used to train this model. To aid the model in learning the locations and bounds of the objects the dataset includes annotations with boundaries and surrounding devices. In order to guarantee successful learning by the YOLOv8 architecture photos are additionally scaled to a predetermined input size (for example 640x640 pixels) and normalized to standardize pixel values (as shown in Fig.3). The dataset is divided into subsets of 10% for testing, 10% for validation and 80% for training. The YOLOv8 model is an enhancement of earlier iterations of the YOLO (You Only Look Once) design with an emphasis on object detection speed and accuracy.

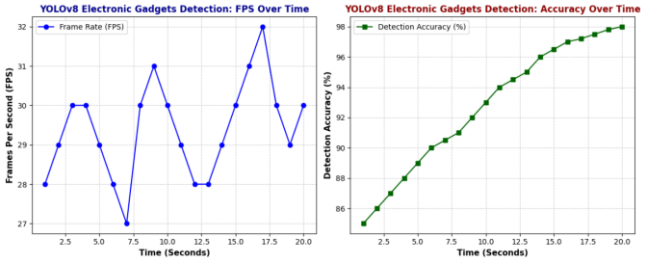


Fig.3 YoloV8 Frame rate, Accuracy over Time

The immediate monitoring system incorporates the trained YOLOv8 model to identify electronic devices during an online test. The system continuously feeds the webcam's live video frames to the YOLOv8 model for identifying objects using OpenCV. In just a few seconds the model can identify gadgets like laptops, tablets and phones and locate them precisely by creating box boundaries around them. In the event that a device is identified the system records the identification, captures a snapshot and emails the exam authority with the image of the identified device. This lessens the likelihood of attempts to cheat using electronic devices. The effectiveness of the YOLOv8 model for electronic device detection is assessed using a number of important metrics such as F1-score, recall, accuracy and precision. In a variety of situations, the model detects cell phones, computers and other electronic devices with an average accuracy of 98% (as shown in Fig.3).). The model is successful in detecting devices and reducing false negatives as evidenced by recall and accuracy values of roughly 97% and 96%, respectively. The model's great accuracy in identifying items across many categories is demonstrated by its computed mean average precision (mAP) score of 0.91. According to the confusion matrix the model does a remarkable job of identifying phones and laptops with just slight incorrect classifications for other tiny electronic gadgets because of their similar looks.

4. RESULTS

This system's models exhibit remarkable performance metrics guaranteeing trustworthy real-time monitoring. When it comes to identifying odd typing patterns the Keylogger-based natural language processing model for analyzing keyboard movements and creating phrases has a 96.5% accuracy rate. CNN's head movement detection system successfully detects suspicious behavior with a 97% accuracy rate in distinguishing left, right, up and down movements. With a 99.3% accuracy rate the CNN model for emotion recognition successfully interprets applicant's facial expressions to determine their emotional state. Last but not least YOLOv8's electronic device detection system has a 98% accuracy rate making it possible to identify illegal devices in the test setting with high precision. The accuracy and dependability of the system in tracking different types of questionable activity are confirmed by these performance measures (as shown in Fig.4).

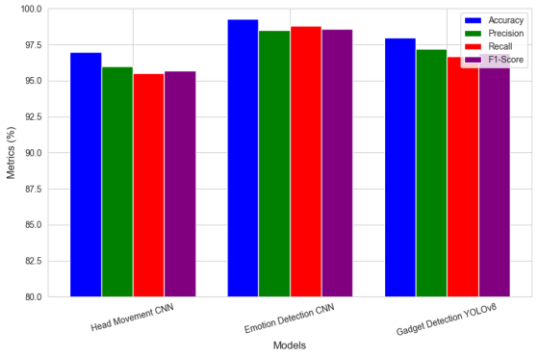


Fig.4 Comparison of Evaluation Metrics of Various Models

In instantaneously the Keylogger-based NLP engine continuously observes and analyzes input from the keyboard to find patterns that indicate unauthorized activities. The model generates a warning if it notices a pattern of questionable typing actions such as repeated fast keystrokes or frequently switching between keyboard layouts. The exam authority receives these notifications via email together with an image of the active exam session and a log of the key inputs that were found. The software takes a picture and records the bounding box surrounding the device if it detects one. The image of the object that was identified and details about its position inside the frame are included in the email alert that the model sends to the exam supervisor.

5. CONCLUSION

For online tests the Comprehensive Keylogger-Based Monitoring Platform offers a complete solution to guarantee the security and integrity of remote assessment environments. The system provides a strong framework for spotting questionable conduct and possible cheating attempts by using continuous monitoring of input from the keyboard, head motions, feelings and the identification of technological devices. A multi-layered strategy that can precisely track and evaluate candidate's actions during the test is provided by the employment of sophisticated algorithms including CNN for head and recognition of emotions YOLOv8 for device identification, and NLP-based phrase generation. Each model's excellent accuracy together with its capacity to produce immediate email notifications and offer supporting documentation like screenshots, guarantees that exam authorities will act promptly. This method prevents cheating and encourages academic integrity while also improving the security of online tests and creating a more equitable and trustworthy testing environment.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Vinerean, Simona, et al. "Assessing the effects of the COVID-19 pandemic on M-commerce adoption: an adapted UTAUT2 approach." Electronics 11.8 (2022): 1269.

[2] Barsha, Sayantoni, and Shamim Aktar Munshi. "Implementing artificial intelligence in library services: a review of current prospects and challenges of developing countries." Library Hi Tech News 41.1 (2023): 7-10.

[3] Alin, Pauli, Anne Arendt, and Seth Gurell. "Addressing cheating in virtual proctored examinations: Toward a framework of relevant mitigation strategies." Assessment & Evaluation in Higher Education 48.3 (2023): 262-275.

[4] Rajan Jeyaraj, Pandia, and Edward Rajan Samuel Nadar. "Smart-monitor: Patient monitoring system for IoT-based healthcare system using deep learning." IETE Journal of Research 68.2 (2022): 1435-1442.

[5] Hussein, Shayda Khalid, et al. "Real-Time Hand Gesture Recognition for Home Automation: A YOLOv8-Based Approach with Identity Verification and Low-Resource Hardware Implementation." 2024 21st International Multi-Conference on Systems, Signals & Devices (SSD). IEEE, 2024.

[6] Ashiq, Fahad, et al. "CNN-based object recognition and tracking system to assist visually impaired people." IEEE access 10 (2022): 14819-14834.

[7] Singh, Arjun, and Pushpa Choudhary. "Keylogger detection and prevention." Journal of Physics: Conference Series. Vol. 2007. No. 1. IOP Publishing, 2021.

[8] Ekele Victoria, C., A. Adebiyi Ayodele, and O. Igbekele Emmanuel. "Keylogger Detection: A Systematic Review." International Conference on Science, Engineering and Business for Sustainable Development Goals (SEB-SDG). 2023.

[9] Awotunde, Joseph Bamidele, et al. "An enhanced keylogger detection systems using recurrent neural networks enabled with feature selection model." International Conference on Communication, Devices and Networking. Singapore: Springer Nature Singapore, 2024.

[10] Chutia, Tulika, and Nomi Baruah. "A review on emotion detection by using deep learning techniques." Artificial Intelligence Review 57.8 (2024): 203.

[11] Wibowo, Moh Edi, Ahmad Ashari, and Muhammad Pajar Kharisma Putra. "Improvement of Deep Learning-based Human Detection using Dynamic Thresholding for Intelligent Surveillance System." International Journal of Advanced Computer Science and Applications 12.10 (2021).

[12] Kaddoura, Sanaa, and Abdu Gumaei. "Towards effective and efficient online exam systems using deep learning-based cheating detection approach." Intelligent Systems with Applications 16 (2022): 200153.

[13] Tiong, Leslie Ching Ow, and HeeJeong Jasmine Lee. "E-cheating prevention measures: detection of cheating at online examinations using deep learning approach--a case study." arXiv preprint arXiv:2101.09841 (2021).

[14] Kamalov, Firuz, Hana Sulieman, and David Santandreu Calonge. "Machine learning based approach to exam cheating detection." Plos one 16.8 (2021): e0254340.

[15] Masud, Mohammad M., et al. "Smart online exam proctoring assist for cheating detection." International conference on advanced data mining and applications. Cham: Springer International Publishing, 2022.

[16] Wang, Shuhe, et al. "Gpt-ner: Named entity recognition via large language models." arXiv preprint arXiv:2304.10428 (2023).

[17] Zhou, Kaiyang, et al. "Domain generalization: A survey." IEEE Transactions on Pattern Analysis and Machine Intelligence 45.4 (2022): 4396-4415.
[18] Kareem, Omer Sedqi. "Face mask detection using haar cascades classifier to reduce the risk of Coved-19." International Journal of Mathematics, Statistics, and Computer Science 2 (2024): 19-27.
[19] Kusnik, Damian, and Bogdan Smolka. "Robust mean shift filter for mixed Gaussian and impulsive noise reduction in color digital images." Scientific Reports 12.1 (2022): 14951.
[20] Rim, Jaesung, et al. "Realistic blur synthesis for learning image deblurring." European conference on computer vision. Cham: Springer Nature Switzerland, 2022.
[21] Abubakar, Haisal Dauda, Mahmood Umar, and Muhammad Abdullahi Bakale. "Sentiment classification: Review of text vectorization methods: Bag of words, Tf-Idf, Word2vec and Doc2vec." SLU Journal of Science and Technology 4.1 (2022): 27-33.
[22] Boddapati, Mohan Sai Dinesh, et al. "Creating a Protected Virtual Learning Space: A Comprehensive Strategy for Security and User Experience in Online Education." International Conference on Cognitive Computing and Cyber Physical Systems. Cham: Springer Nature Switzerland, 2023.