ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN: 2454-9940 www.ijasem.org
Vol 20, Issuse.2 April 2026

Exam Watch AI: Real-Time Intelligent Online Exam Proctoring System

[1]Nimmakayala Pravalika,[2] Ch Mahesh

[1]M.Tech Scholar,Dept. of CSE, Malla Reddy Technical Campus,Malla Reddy Vishwavidyapeeth, Maisammaguda,Hyderabad, Telangana 5001400, India, npravalika999@gmail.com
[2]Assistant Professor, Dept. of CSE, Malla Reddy Technical Campus, Malla Reddy Vishwavidyapeeth, Maisammaguda, Hyderabad, Telangana 500100, India, chekurimahesh364@gmail.com

Abstract

Online exam circumstances make it harder to identify deviant activity due to sophisticated monitoring settings and ever-changing student movements. The ability to identify suspicious actions in real-time relies on precise behavioral modeling and sophisticated analysis of video feeds. Using computer vision and machine learning, this work presents an AI-powered real-time proctoring system that can automatically monitor examinee behavior. The proposed system primarily operates via three stages: anomaly identification, behavioral state analysis, and face detection. Initially, student faces are identified using the Haar Cascade approach in order to enable continuous presence verification. Next, we'll use a trained neural network to analyze the student's facial expressions, movement patterns, and other behavioral markers to see whether they're behaving suspiciously. Finally, anomaly detection uses a Gaussian distribution model to evaluate the frequency of suspicious events within a certain time range. The program may identify instances of potential malpractice when it finds abnormal patterns that exceed certain parameters. To enhance automated invigilation, the proposed technique aims to increase online exam reliability, accuracy, and fairness while minimizing human supervision work. This technology provides a safe alternative for remote assessment by enabling intelligent decision-making in real-time and continuous monitoring.

Keywords: AI-Based Proctoring; Hand Movement Detection; Deep Learning; Media Pipe; Eye Movement Detection

Introduction

The field of computer vision and understanding human behavior has lately attracted a lot of attention since it is one of the most challenging, complicated, and distinctive ones.just one. Investigators engaging in overly tedious and, at times, ineffectual supervised strict monitoring is the standard practice for test hall invigilation. A computer vision videotape analytics operation that analyses surveillance footage of a crowded test hall will have its biggest challenge in the massive quantity of background processing that is necessary to develop it. In the name of purportedly unlawful conditioning, there are often victims whose faces must be acknowledged, respected, and concealed. The design of the testing facility necessitates near-real-time execution of all findings and subsequent processing. The computer vision operation becomes much more intricate as a result of this. Additional obstacles to the effectiveness of such an intelligent surveillance system include occlusion and image depth. As an example, it's quite likely that examinees positioned at the far end of the camera will manage to avoid being caught. Candidates should be permitted to make little, natural movements, like their hands moving while writing, and these should be ignored, even if the stir is the first stage in training. We can't promise that what people say will be true. Instead of depending on one state (frame) to identify an anomalous geste, it would be great if the software system depends on many countries. Right now, the hardest part is coming up with an innovative algorithm that relies on literacy. Face discovery (2) and recognition (3-4) will be vital in this process for identifying and fete-ing examinees against a pre-populated database of campaigners. To

ISSN: 2454-9940 www.ijasem.org

Vol 20, Issuse.2 April 2026

ensure that examinees do not participate in counterfeiting, facial recognition technology must be very precise and dependable. Locating and recognizing faces is the first and most crucial step. In the realm of security, facial recognition technology is crucial. Many people think that a person's facial features and stride are the most crucial biometric traits for identifying them in video surveillance systems. Facial expressions may be identified by observing changes in the displaced facial traits (5). Winking or denying a headshake are two examples of the many ways some people use their faces to change the meaning of what they say. Raw facial expressions make it difficult for software to tell if an update is being made to factual information or whether these feelings were just informal. It would be perfect if, even in a very crowded exam room, a computer vision videotape logical operation could be designed to keep an eye on every student in real time. Section 2 will provide a brief overview of the scientific basis of the habitual ways, with brief references to the connected material, due to space limits. Section 3 lays out the recommended procedure.

New opportunities for developing intelligent monitoring systems that can assess human behavior in real-time have emerged as a result of advancements in computer vision and artificial intelligence. The identification of anomalous activity in video surveillance has been a hot issue because to its extensive application in several sectors such as public safety, healthcare, traffic management, workplace monitoring, and educational institutions. One of these areas that presents a particularly significant challenge is exam monitoring. Exams must be fair, disciplined, and honest if applicants are to be monitored constantly at all times since even little variations in behavior might indicate suspected cheating. Because of their reliance on human error and inability to manage large exams, manual invigilation methods have significant drawbacks. So, it's necessary to have smart technology that can automatically detect and classify suspicious conduct in exam rooms. Problems arise while trying to discover anomalies in videos because real-world conditions are dynamic and unpredictable. What may be seen as suspicious in one context may be perfectly appropriate in another. While a test taker who glances about the room a lot can be seen as suspicious in a more formal setting, it might be considered as normal behavior in a more casual one. Combining probabilistic modeling with pattern recognition and robust feature extraction is necessary for algorithms to distinguish between normal and abnormal behavior. Even more complicated are factors like as lighting, occlusions, moving objects, and individual heterogeneity in the appearance of anomalies. Therefore, anomaly detection ranks high among the most challenging and open-ended issues in video analysis.

Literature Survey

The Haar Cascade classifier, which is now one of the most popular approaches to face recognition, was first created for use in real-time object detection in this study. Using a combination of Haar-like features and an integrated image method allows for computationally efficient detection. Despite this method's roots in face identification, it has found extensive utility in activity recognition systems, particularly in surveillance applications. When integrated with neural networks, Haar features enhance video surveillance by tracking motion patterns in individual frames and accurately localizing faces, making it simpler to detect suspicious behaviors.

This article gives a summary of anomaly detection methods and stresses the significance of discovering unusual patterns in time series data. It shows how hard it is to represent human activities since they are so situational and changeable. The authors provide a strategy for detecting out-of-the-ordinary or normal behavior by going over statistical, proximity-based, and machine learning methods. This paper establishes the framework for using anomaly detection techniques in limited space monitoring systems, surveillance, and fraud detection.
We provide deep learning techniques for anomaly identification in videos in this study. The authors provide a method for deep multiple instance ranking using anomaly scores that can be applied to video segments. Unlike previous methods, our system can detect unique and minor anomalies in long surveillance records. Due to the high importance of uncommon but substantial abnormalities, their work is especially beneficial in classroom and test hall surveillance. When compared to feature-based techniques that are manually built, neural networks significantly outperform them when it comes to spotting suspicious behavior.
This study demonstrates the use of Gaussian Mixture Models (GMMs) for statistical modeling and outlier detection. We discover that GMMs adequately

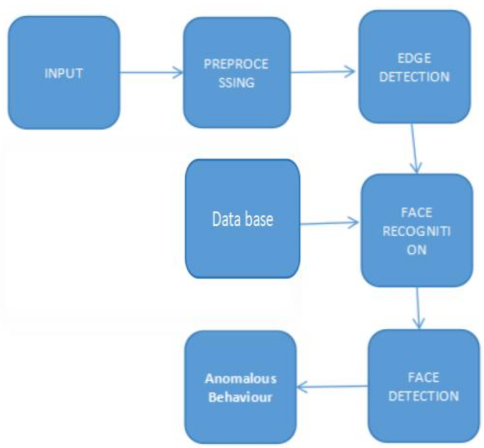
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN: 2454-9940 www.ijasem.org

Vol 20, Issuse.2 April 2026

represent the normal distribution of probability for behavioral patterns while also being able to detect outliers that may indicate an abnormality. The statistical foundations provided by Gaussian models are crucial to systems that depend on probabilistic reasoning, including those that scan examination contexts for repeated suspicious events. The results of this research suggest that hybrid detection systems that use machine learning in conjunction with generalized linear model (Gaussian) models may achieve better results.

Methodology



Proposed diagram

Gather raw visual data (the foundation upon which everything else is built). It includes a video stream or series of frames captured by security cameras set up to watch examinees. Metadata such as the camera's ID, time stamp, frame rate, and resolution may be included as well.

Work carried out at this location

Record video at a predetermined frame rate, say, 10 to 30 frames per second. To lessen the burden, you may choose to use lightweight frame sampling, which involves analyzing every third frame, for example. Important things to think about and ways to put them into practice: If you want to avoid occlusion while detecting faces, it's best to use a frontal or slightly above view, therefore be cautious with your camera positioning and quality. Very high resolution adds computation, although 720p is usually sufficient. Maintain a constant frame rate and synchronize timestamps. To prevent frame drops, use a buffer. To facilitate offline model training or forensic examination, it is possible to record and save raw streams; however, stringent privacy constraints should be implemented.
This block passes on to the Preprocessing block the ordered frames (images) that include timestamps and camera identification. Occlusions (hands/objects), motion blur, several subjects in the frame, and changing lighting conditions are common problems. Precautions: a well-positioned camera, infrared lighting for dim conditions, and accurate shutter and exposure controls.
Initial Steps The goal is to enhance the reliability and efficiency of subsequent detection and classification by cleaning and standardizing input frames. The inputs for edge detection, face detectors, and classifiers are normalized from noisy heterogeneous video frames by preprocessing.

Standard procedures and algorithms:

One, color space conversion: numerous detectors may be converted to grayscale (which simplifies processing), and an RGB copy can be kept for appearance characteristics if desired. Decrease sensor noise by using a median filter or a Gaussian blur. Usually, kernel sizes are 3×3 or 5×5. Thirdly, histogram equalization and illumination correction: Systems such as CLAHE or basic histogram equalization enhance contrast in all lighting scenarios. Scaling and normalization: Set the dimensions of frames or ROIs to a constant value, such as 160×160 or 224×224 for NN inputs, then scale the pixel values to either a range of 0 to 1, or a meaning-zero unit variance. Frame differencing or background subtraction: You may use this as an optional motion emphasis technique (like MOG2 or basic frame-diff) or you can utilize Otsu's approach to automatically threshold when you derive binary foreground masks. Data enhancement (during training): data flipping, rotation, and modest brightness/contrast adjustments to strengthen models. Realistic recommendations for parameters The median filter radius is 1-2; the Gaussian blur is around 0.7-1.5. Resize such that it fits the input size that downstream models demand (without stretching the aspect ratio; pad as necessary). To prevent false

thresholds while using Otsu for thresholding, it is recommended to work on a smoothed picture.
Results: Frames that have been cleaned up, binary foreground masks, and standardized ROIs that are sent to either Edge Detection or face detectors straightaway. The loss of fine face information due to over-smoothing or uneven NN performance due to insufficient normalization are two common problems. Preprocessing should always be double-checked using a set of frames that reflect the most extreme lighting and occlusion scenarios.
detecting the edge
The goal is to locate face contours, identify the outlines of heads and hands, and provide features for landmark extraction by extracting important structural characteristics (edges and contours) from preprocessed frames. Feature extraction and downstream face-alignment are often assisted by edge maps.

Regularly used algorithms: Two-threshold hysteresis (low/high thresholds) to generate clean edge maps is used by the most popular edge detector, Canny. Sobel and Prewitt: good supplementary characteristics for gradient magnitude and direction. Laplacian of Gaussian: used for edge detection on several scales.
Reasons why edge detection is useful for this task: When texture and contrast are poor, it helps to distinguish the edges of faces. By limiting potential areas, it strengthens landmark detection (eyes, nose, mouth). Makes it easier to separate motion and gestures (such as a head-turn outline from the backdrop).
Adjustments and settings: Adaptively setting low_thresh and high_thresh (for example, according to median pixel intensity) or computing them per-scene to deal with illumination fluctuation are two important considerations for Canny. For 8-bit pictures, the usual approach is to set low = 0.66 * median and high = 1.33 * median, or to put low > 50 and high > 150. To decrease the number of false positive edges, pre-smooth using a Gaussian kernel that is calibrated to the noise level. Output: Orientations of gradients and binary or weighted edge maps. Some behavior heuristics, such as the quick change in head shape over time, might benefit from these as lightweight characteristics, and they also help with facial landmarking and guiding face alignment.
Common concerns and their remedies include: edge noise (which can be reduced using Gaussian smoothing), background texture (which may be addressed using ROI from face detector to focus), and lighting sensitivity (which can be addressed by using contrast normalization before edge detection).

Recognition of Faces

Goal: Use a database of registered examinees to confirm or deny the identification of the observed face. This block allows for the monitoring of behavior over time per individual by associating visual evidence with that person.
Contrast: face detection seeks for faces in a picture (where). A person's identity may be determined using facial recognition technology. Prior to recognition, several pipelines do detection. The placement of these blocks in the figure implies that both detection and recognition are present; in actuality, implement detection first, then recognition.

Techniques and methods:

Pretrained embedding networks (FaceNet, ArcFace, or ResNet from dlib) are used to transform a face crop into a fixed-length vector in modern solutions. From there, recognition is reduced to nearest-neighbor search (cosine distance) or a trained classifier (SVM, softmax). Compared to embedding-based algorithms, previous approaches such as template matching, eigenfaces, and Fisherfaces are not as robust. Instructions for execution and settings: To ensure uniform embeddings, align the face (using facial landmarks) before inserting the device. Set a cosine similarity criterion greater than 0.6 or adjust it based on the validation set as an embedding threshold for identity matching. As a general rule, it's best to save many embeddings of each user (with varying lighting and poses) and then use the minimal distance to make a decision. Interaction with the database: Check the database for registered templates and update the logs when recognition happens. Results: "unknown" or "identity" label, matching score, and confidence in the identity. For round-the-clock surveillance, feed the Face Detection/Tracking block your identify and ROI. For use in a test setting: Associate suspicious occurrences with a particular examinee, resulting in notifications that are unique to that individual. Helps compile data on a per-person basis (how often this individual displayed questionable behavior).Problems and solutions: Low resolution and pose variation (in-person masks, head tilt) lower recognition accuracy; illumination and low resolution employ enhancement and higher-resolution capture for enrollment; and low resolution and pose variation diminish multi-view enrollments.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

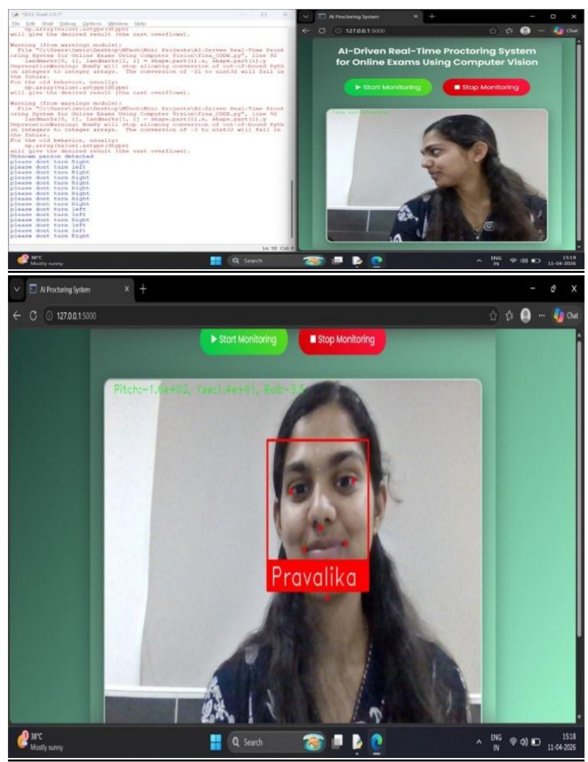
ISSN: 2454-9940 www.ijasem.org

Vol 20, Issuse.2 April 2026

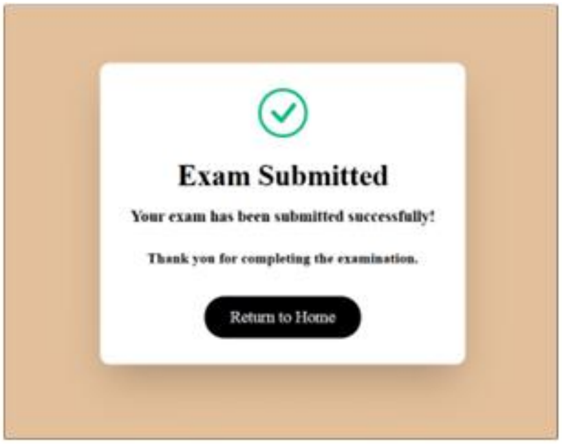
Face Detection

In order to study behavior across time, it is necessary to continuously find and monitor facial areas across frames. In this case, the block is used for runtime bounding-box extraction and tracking after recognition/identification, rather than detection, which is the first step in certain pipelines. Essential duties within this section: First, identify faces in each frame using contemporary CNN-based detectors (e.g., MTCNN, RetinaFace) or Haar cascade (Viola-Jones). While Haar is quick and works well with frontal faces, MTCNN and RetinaFace are better at handling lighting and poses. Alignment of the face and landmark detection: Standardize the stance and get ready for expression or recognition analysis by detecting the eyes, nose, and mouth. Software: MediaPipe Face Mesh, dlib. In order to stabilize analysis and decrease computation, run a tracker (such as KCF, CSRT, MOSSE, or correlation trackers) once a bounding box has been discovered. This will retain the face in memory between detection runs. Extracting features for behavior analysis: calculate yaw, pitch, and roll vectors for the head, as well as the ratio of open to closed eyes, gaze proxy, mouth openness, and movement velocity. This data is sent into the classifier that looks for suspicious states. Realistic guidelines and pointers: To decrease the number of false positives produced by HaarCascade, you may adjust scaleFactor (for example, 1.1-1.2), minNeighbors (3-6), and minSize. • Use a detector at regular intervals (every N frames, for example, 5-10 frames) and trackers in between. Re-identification may be handled using the recognition block, and per-face IDs should be maintained. Results: Bounding boxes for each monitored person's face each frame, landmark locations, head-pose/gaze estimations, and a permanent face ID. After that, the Anomalous Behaviour determination is informed by the Suspicious State Detection (neural network). Most often encountered issues include tracking drift, which may be resolved by starting with a new detection, false-positive non-face ROIs, which can be confirmed using recognition or a tiny NN classifier, and multiple faces crossing, which can be handled with strong re-id logic.

Results



Warning system



Auto Submission

Conclusion

In this work, we used examinee monitoring during tests to approach the challenging topic of detecting anomalous behavior in complex contexts. The scenario becomes worse when you include in the intricacies of abnormal movements, all the moving parts in exam rooms, and the need to distinguish between regular and suspicious movements. Using a normal distribution model for anomaly identification, neural networks for suspicious state detection, and Haar cascade classifiers for face recognition, we constructed a prototype of a monitoring system to circumvent these issues. We started by making sure the machine was focusing on the correct places—the examinees' faces—using Haar cascade classifiers for face detection. Consistent monitoring of students in a variety of seating arrangements and environmental settings was made feasible by this method, which also reduced background clutter. Adopting the Haar-based approach proved that it was appropriate for real-time applications because to its processing efficiency and durability for frontal face detection. In the second stage, we trained neural networks to detect suspicious situations, which significantly enhanced behavior analysis by revealing patterns that human rule-makers may overlook. Neural networks have the potential to distinguish between regular head and eye movements and those that may indicate suspect conduct, such as staring at the test paper excessively, keeping an unusual gaze direction, or displaying strange facial alignments. This machine learning-based approach allowed the system to be both adaptable and scalable by retraining the model with new data to improve its accuracy and generalization capabilities over time.

REFERENCES:

[1]. chandra Sekar P. and et.al(2022) "An Energy Efficient Architecture for Furnace Monitor and Control in Foundry Based on Industry 4.0 Using IoT", Scientific Computing for Internet of Health Informatics Things ,Volume 2022 ,Article ID 1128717.

[2]. https://doi.org/10.1155/2022/1128717A. S.Sakena Benazer, M. Sheik Dawood, "Design of Rectenna for wireless sensor Networks" Elsevier, Materials Today: Proceedings, 2021, ISSN 2912-2915.

[3]. Sakena Benazer. S, Sulthana, T. Anish and Ulageshini, S. "Analysis of Pandemic COVID-19 Impact in India" Journal of Science Technology and Research (JSTAR), Volume No.1, Issue No.1 (2020).

[4]. R. Uma Maheshwari, B. Paulchamy, Arun M, Vairaprakash Selvaraj, Dr. N. Naga Saranya and Dr . Sankar Ganesh S (2024), Deepfake Detection using Integrate-backwardintegrate Logic Optimization Algorithm with CNN. IJEER 12(2),

[5]. Maheshwari, R.U., Kumarganesh, S., K V M, S. et al. Advanced Plasmonic Resonance-enhanced Biosensor for Comprehensive Real-time Detection and Analysis of Deepfake Content. Plasmonics (2024). https://doi.org/10.1007/s11468-024-02407-0

[6]. V.Malathy, M.Anand, S.M.Kamali & G.Umadevi, (2020), "Parallel pipelined multi radix variable length fast fourier transform Architecture", ISSN- 2394-5125 Vol 7, Issue 6, 2020, https://www.researchgate.net/publication/341932765.

[7]. MOORTHY, V. P., SUBRAMANIAM, K., KAMALI, S. M., & GEETHA, V. (2023). "INVESTIGATION ON DIELECTRIC PROPERTIES OF BIO-TRANSOL FOR TRANSFORMER INSULATION APPLICATIONS". In. Journal of Environmental Protection and Ecology, Vol. 24, Issue 7, pp. 2395–2403. https://scibulcom.net/en/article/5jSd7tSqDnunHUDh5tdY.

[8]. Prasanna Moorthy Venugopal., Kamali Samudram Manickam., Malathy Vanniappan., & Gowri Sankar Periyannan Annadurai., (2024), "Simulation and Electrification of aConventional Vehicle into a Battery Electric Vehicle for Sustainability Mobility" International Journal of Energy Sources, Part A: Recovery, Utilization, and Environmental Effects,Vol.46,No.1, pp. 6649-6661. doi: 10.1080/15567036.2024.2353192

[9]. P. Babu, M. I. Habelalmateen, R. Srikanteswara, R. A. Reddy and N. Purushotham, "Wafer Surface Semiconductor Defect Classification Using Convolution Neural Network Based Improved Faster R-CNN," 2024 Second International Conference on Data Science and Information System (ICDSIS), Hassan, India, 2024, pp. 1-4, doi: 10.1109/ICDSIS61070.2024.1059430.

[10]. 5.Sakena Benazer. S, Sulthana, T. Anish and Ulageshini, S. "Analysis of Pandemic COVID-19 Impact in India" Journal of Science Technology and Research (JSTAR), Volume No.1, Issue No.1 (2020).

[11]. Hariprasath Manoharan, Sulaima Lebbe Abdul Haleem, S. Shitharth, Pravin R. Kshirsaga, Vineet Tirth, M. Thangamani, Radha Raman Chandan, A machine learning algorithm for classification of mental tasks, Computers and Electrical Engineering, Vol. 99, No.107785, 2022

[12]. V.Malathy, M.Anand, S.M.Kamali & G.Umadevi, (2020), "Parallel pipelined multi radix variable length fast fourier transform Architecture", ISSN- 2394-5125 Vol 7, Issue 6, 2020, https://www.researchgate.net/publication/341932765.