International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

Human Monitoring and Detection in Different Environments: Review

Boya Durgabhavani[1]
Research Scholar, JAIN(Deemed-to-be University), Bengaluru, Karnataka, India.
durgabhavani.b@jainuniversity.ac.in

Dr. Sagar Gulati[2]
School of CS and IT, JAIN(Deemed-to-be University), Bengaluru, Karnataka, India.
Sagar.gulati@jainuniversity.ac.in

Abstract **- Human Monitoring and detection have appeared as a recognizable research field owing to the advancement of computer vision, machine learning, and deep learning. All three technologies allow the analysis of human existence, actions, and behavior. This work offers a structured review of the existing literature on human monitoring and detection, specifically in academic and office settings. It encompasses the concepts of attendance, engagement analysis, smoking detection, and occupancy analysis in the different environments.**

The most notable advancement mentioned in the literature review is the evolution of face processing methods from conventional manual image processing algorithms to deep learning-based approaches such as YOLO algorithms, CNN, multimodal frameworks, among others. Other aspects covered are sensor-based or edge-supported face processing frameworks proper fit for locations where technical resources are small. However, despite the various advances in face processing technologies, the current gaps in environmental changes, computational complexity, data availability, as well as privacy issues associated with face processing remain. Future lines should involve multimodal frameworks, edge-AI models, privacy-preserving frameworks, behavioral analysis, diverse datasets, among others. Moreover, combining temporal data with visual indicators could offer a more effective means for comprehending human behavior and actions in various contexts.

Keywords - Human monitoring, Human detection, smoking monitoring, Employee monitoring, activities, behavior.

I. INTRODUCTION

- Human monitoring and detection systems stand as a pivotal and rapidly evolving area of research covering computer vision, machine learning, and deep learning [1]. It becomes essential in educational institutions, offices, hospitals, and public spaces due to increasing needs in safety, security, and intelligent automation [2]. With the significant advancement of scientific and technological innovation in various sectors, the demand for monitoring systems is growing, as these systems help in increasing security and productivity [3]. The core focus is on Human Activity Recognition (HAR), which seeks to identify, analyze, and categorize human body movements, postures, and behaviors across various real-world visual and sensor-based scenarios [1] [4]. To detect the presence of individual in the open spaces is commonly use multiple Technologies including sound-based sensors, image-based human recognition, infrared detection, chemical sensing and others [5]. Therefore, an automatic video

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

surveillance system is required to identify, track, and analyze the individual's behaviors in the videos footage [6]. Person detection in security footage used to be done manually. However, there were a significant number problems with that strategy. First, it was too much for one individual to watch the video in an effort to detect human presence continuously. As such, several individuals was needed to work on a rotational shift wise. Second, because no one can sit still and watch the video for several hours at a time, human negligence is another issue that cause important portions of the video to be missed [7] [8] [9].

A. Types of Detection Implemented in Academic Settings

Today, most colleges utilize rapid advancement in technology to get the job done. Automatic systems can manage tasks that was traditionally performed by humans, such as automatic attendance system in schools and colleges [10]. Human monitoring and detection in college classrooms primarily relies on deep neural networks and visual computing systems to capture and classify physiological states, attention levels, and specific student actions.

1) Student Attendance Monitoring

K AI Hajji et.al. implemented Automated Attendance Monitoring System in colleges, aiming to modernise the outdated and inefficient traditional methods of recording student presence, such as calling roll or using paper sign-in sheets [11]. Facial identification attendance features offers high accuracy and enables rapid identification of individuals within seconds. In 2025, Sakthikumar et al. introduced an AI-powered intelligent attendance tracking system using YOLOv8. They conducted a comparative analysis between YOLOv8, RCNN, and Haar Cascade, where YOLOv8 outperformed the others, achieving an impressive accuracy of 97.5% [12]. In 2023, Deepa et al. implemented a computer vision-driven intelligent use for monitoring student in classroom. A facial recognition accuracy of 74% is achieved using YOLOv7 on a dataset of 224 images. Finally, the implementation of the proposed framework follows a three-fold stage process involving the following:

1. For the classroom monitoring API service, Firebase is employed as the backend.
2. Integration of the face recognition system is achieved using YOLOv7, and
3. Developing a system for identifying individual students [13].

Raj et al. proposed the design of a Facial Recognition-Based Student Attendance System, with steps such as database creation, image analysis, facial localization, feature representation and facial identification, redundancy removal, and report generation. For better image processing, the researchers used GANs, and the process of detecting faces involved 70 facial landmarks and Haar classifiers. Gabor filters were used at different angles to ensure proper extraction of features. For face recognition, the proposed system used three algorithms: SVM, CNN, and KNN, with the best accuracy of 92% by KNN, while the accuracy of CNN was less but more computationally efficient compared to the inefficient process of SVM [14]. Rajesh et al. proposed a system using a self-built dataset. The process of the designed framework for attendance generation involves acquiring the input image, face localization, facial region segmentation, visual feature extraction, face recognition with the trained database, and report generation. In the "Feature Extraction" stage of the proposed system, the designers used

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

sophisticated techniques, including the use of CNN, to meticulously identify the minute details and features of faces in the different images [15].

2) **Student Engagement Monitoring:** Student engagement encompasses the time and psychological effort a student devotes to and is committed to their learning experience and the learning environment as a whole. According to [16], student performance is based on classroom engagement. In 2025, Habib et al. developed a system for visual cues and pose estimations for the automatic categorization of levels of engagement-Engaged, Not Engaged, or Partially Engaged [21]. They used learning-based algorithms and visual data processing technology. The utilization of YOLO is basically for object recognition; ML models can take data collected through imaging sensors such as cameras to monitor the physical features of students and their behaviors. In 2024, Amish Abraham et al. proposed a machine learning-based engagement monitoring system. The system was designed with graphical representations of Eye Aspect Ratio (EAR) and yawn values for the evaluation of students' attentiveness related to the status of being focused or distracted [17].

Ritu Dudhmal et al. suggested a AI driven approaches for monitoring students in online classes. They have detected 15 different student behavior patterns in online classes and used a Convolutional-driven learning model (CNN) to recognize instances of distractions [18]. Sakib et al. promoted a deep learning-based technique for classroom activity grouping to measure student engagement based on reading, writing, and side-talking activities. In their research, they incorporated various models including Logistic Regression (LR), Random Forest Classifier (RFC), Support Vector Machine (SVM), and XGBoost (XGB) [19]. Ami et al. suggested a method for student pose estimation and attention assessment level prediction using DeepLabCut and SVM classifiers. Their work uses DeepLabCut and SVM to predict student attentiveness level into fully engaged, partly engaged, and not engaged categories [20].

3) **Student Behavior:** Analysis of student behavior has a significant role in not only measuring the performance of teaching but also evaluating student development over an extended period[22]. It is employed extensively for increasing the learner experience, increasing student performance, and managing classroom behavior efficiently [23]. In 2023, Akash et al. proposed a machine learning technique for real-time student behavior analysis, emphasizing the prediction of eye detection, yawning, and head position. The technique has resulted in a prediction accuracy of 83% for yawning, 73% for the detection of eyes, and 82% for head position estimation [24]. In 2024, Hamidi et al. implemented a real-time technique for drowsiness prediction of students incorporating deep learning methodologies with YOLOv8. It recognizes the physical manifestation of drowsiness, such as eyes shutting and yawning, which is likely to happen when students lose attention. It aims to detect whether a student is awake or drowsy [25].
In 2024, an intelligent recognition algorithm for college students' classroom behavior classification was introduced by Cui et al., who used an enhanced version of the SSD framework. The standard SSD framework had some drawbacks in recognizing the four targeted behaviors – listening, writing, playing, and sleeping. The drawbacks were overcome by adding an attention mechanism to the VGG-16 framework, which was based

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

on a ResNet residual network. The purpose was to enhance the capacity to recognize target behavior patterns, particularly when dealing with object obstruction [26]. Santoso et al. proposed a modified version of YOLOv9 incorporating a Bi-Level Routing Attention mechanism for student behavior monitoring in densely populated classroom environments. The system was designed to detect three specific behaviors: hand-raising, reading, and writing, achieving an overall accuracy of 80%. The researchers recommend the subsequent studies that future research should focus on developing more efficient methods to enhance accuracy without compromising detection speed, as well as improving the balance of behavior annotations within the dataset or expanding it to cover a broader range of student behaviors in classroom settings [27].

4) **Occupancy Detection:** In 2022, L Monti et.al developed a edge-enabled learning model for classroom occupancy detection. The researchers used two major classes of datasets CSC and COCO dataset. A learning based algorithm that was pre-trained using the ImageNet dataset, then trained using two specific datasets. A fat client-thin server architecture, in which the prediction was computed directly at the edge [28]. Sha, O et. al. present a student detection and people counting system using the Apriori Algorithm. The Apriori algorithm represent a technique for frequent pattern discovery within a given dataset. It is particularly effective in identifying patterns that frequently appear together. The system struggles with shadow and illumination changes due to its linear body segmentation, resulting in a lower correct detection rate [29].

In 2024, Challa et.al. focused on Real-Time classroom occupancy detection by optimizing HVAC efficiency via Deep Networks. YOLOV4 Object detection model trained to recognize human faces. The paper presents a robust CV pipeline for precision classroom occupancy monitoring [30].

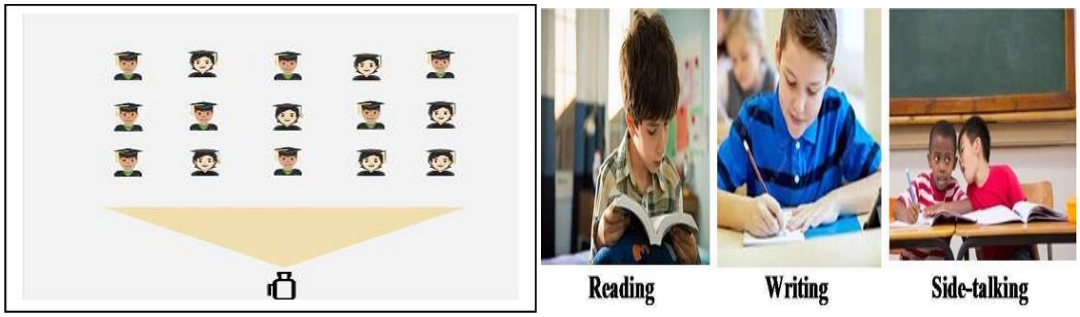


Fig 1: Sample Images

B. Types of Detection Implemented in Office Settings

Monitoring and supervision of employees is essential for sustaining productivity and engagement in the workspace. Attention and engagement of the employees can be reduced while performing operations that are repetitive and boring. Traditional supervision involves a human supervisor monitoring employees' activities; however, constant human intervention in the supervision process is laborious and can be prone to errors, negligence, subjectivity, and bias.[31].

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

1) **Employee Activity monitoring:** In 2023, Harshavardhan et. al. focused on CCTV-based system for monitoring employee activities. By taking into account their posture to ascertain if they are involved in any work-related tasks[31].



Fig 2: An Example of Human Pose [31]

In 2023, Dmitry Nazarov et al. highlighted an Intelligent service for monitoring employee activities by collecting employee computer screenshots. The Intelligent model uses OCR via the Tesseract library to convert screenshots into text. Magick Library pre-processes images. The system focuses on monitoring screen content to determine if computer time is related to work task (keyword matching) [32]. In 2025, Arun Kumar et al. present a Hybrid multimodal AI framework (HMAI-ERA) combining Face Recognition, screen activity monitoring, mouse and keyboard tracking and workplace motion tracking. Classifies employees as “working”, “not working” or “warning (for suspicious activity like repetitive mouse movement) [33]. In 2022, Gopal et al. highlighted a system that monitors remote employees for drowsiness, triggers alarm/ records video evidence upon reaching a threshold, and uses an SMS API to notify the employer. The system achieved a CNN model validation accuracy of (4.27%) [34].

2) **Employee Attendance Monitoring:** In terms of time complexity relative to biometric system, the author proposes an attendance system based on IP and geo-tracking. The implementation utilizes an Arduino connected to reader components for the IP tracker, and features a system of NFC module composed of an IP tracker and RFID module[35]. The results showed that IP and Geo Tracking systems perform superior than biometric system, with accuracies of 81.25% and 79%. In 2022, Memane et al. proposes biometric fingerprint sensor. The system uses a Raspberry Pi as the controller, linked with an Arduino, a fingerprint sensor (specifically the R307 optical fingerprint scanner), and a camera module [36]. In 2023, Pawar et al. propose a system that uses liveness face recognition combined with geofencing technology for automated, two-factor authentication. Faces are detected using OpenCV and a Haar cascade. It will only acknowledge registered employees and ascertain their physical presence at the stipulated venue before recording their attendance.[37].

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

C. Types of Detection Implemented in public Settings

1) **Smoking Monitoring:** Nevertheless, the smoking habit is recognized as one of the world's biggest public health issues in view of its contribution to health problems, early deaths, and economic impacts [38]. In 2024, Li et al. suggested a transfer learning YOLOv8 model. From the outcome of the experiment, it is clear that the new transfer learning method shows a higher level of accuracy, recall rate, and detection time when compared to traditional approaches [39].



Fig 3: Sample Image of smoking detection [39]

In 2023, Jin et al. The suggested approach is a smoking behavior monitoring and identification system using the **YOLOv5** target detection algorithm. The system detects two categories: "xiangyan" (cigarettes) and "fire" [40]. The system, named Smoke Sentry, utilises an enhanced YOLOv9 framework to detect real-time cigarette objects, with a comparative evaluation against YOLOv8 [41].

2) **Occupancy detection:** In 2022, Patwal et al. proposes a crowd counting model based on a CNN using ResNet50 as the backbone architecture. When tested on the Mall dataset, the proposed technique recorded an MAE of 2.4 and an MSE of 8.7 [42]. In 2022, karlita et al. Proposes a Multi Scale CNN (MS-CNN) architecture using a density map as the learning target. The frontend layer (low-level feature extractor) consists of the first 10 layers of VGG16. The backend layer (high-level feature extractor) uses three stacks of the Inception-ResNet-A module (from Inception-ResNet-v2). The research also introduces Pixel-Based MAE and Pixel-Based MSE as new evaluation matrices to better measure the consistency of the resulting density map [43]. Blangiardi et al. presents a feasibility study using Impulse Radio Ultra-wideband (IR-UWB) sensors (Novelda X7) to count people in a static bus environment. our models were tested: Linear Discriminant Analysis (LDA), Random Forests

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

(RF), 1D CNN, and ResNet18. Deep learning models (CNNs/ResNet) performed worse than simpler ML classifiers (LDA/RF) in data-scarce scenarios due to overfitting [44].

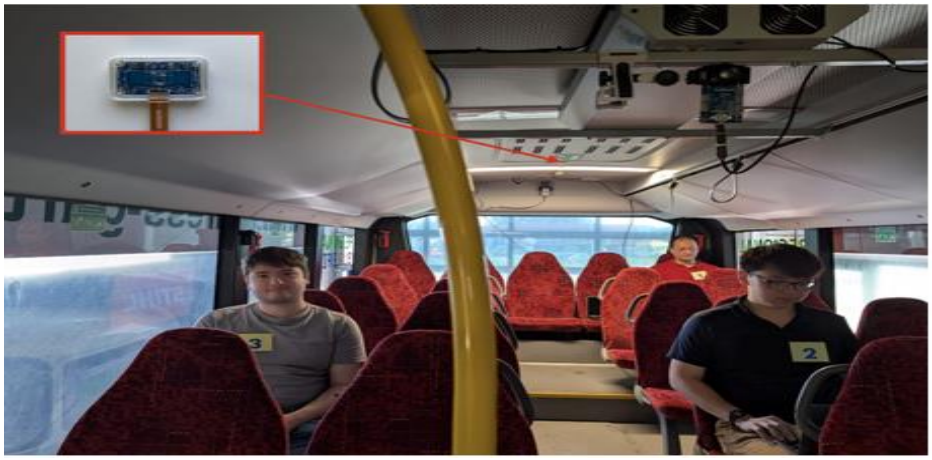


Fig 4: Occupancy detection in Bus[44]

Table 1: Some Common Public Datasets

S. No.	Dataset Name	Application	Key Limitations
1	**COCO**	Person detection; object-based activity detection (e.g., smoking, mobile phone usage)	Imperfect data representation and annotation quality, including boundary inaccuracies, missing labels, and annotation bias.
2	**PASCAL VOC**	Object detection; smoking and mobile phone detection	Limited number of object categories and a relatively small set of annotated images.
3	**UCF101**	Human activity detection (e.g., walking, running, jumping)	Short average clip length (~7.21 seconds), which may be insufficient for modeling complex or long-duration actions.
4	**HMDB51**	Human motion and behavior recognition	High intra-class variability due to videos collected from diverse sources such as movies and online platforms.
5	**NTU RGB+D**	Human activity recognition; posture analysis; fall detection	Data quality issues and noise, with recordings captured under controlled laboratory environments, limiting real-world generalization.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

6	**AVA Dataset**	Spatio-temporal human action detection in videos	Action ambiguity and susceptibility to annotation noise due to fine-grained and overlapping action labels.
7	**EPIC-KITCHENS**	Egocentric human activity detection in indoor and workplace environments	Strong bias toward kitchen-centric activities, limiting applicability to general surveillance scenarios.
8	**FER-2013**	Facial emotion recognition (e.g., happy, sad, angry, fear)	Noisy labels, class imbalance, low-resolution images, and grayscale-only facial representations.
9	**VIRAT Dataset**	Surveillance-based human activity and behavior detection	Low-resolution data and adverse camera viewpoints, affecting reliable action recognition.
10	**PoseTrack**	Multi-person pose tracking and activity analysis	Limited annotations and relatively small dataset size, restricting action diversity and scalability.

PRIVACY PRESERVING AI

Human monitoring and detection in classrooms and offices can raise significant privacy and ethical concerns because of the continuous monitoring conducted by incorporating cameras [45]. In addition, human monitoring and detection systems are capable of monitoring facial information, activities, and individual information. In light of these concerns, Privacy-Preserving Artificial Intelligence (PRAI) approaches are used with the intent of ensuring privacy while carrying out accurate monitoring.

A) Need of Privacy Preservation in Monitoring Systems

Traditional monitoring systems capture and store raw video streams and images, which can lead to [46]:

- Unauthorized Identity detection
- Misuse of recorded videos
- Discomfort among students and employees
- Legal violations

B) Visual Privacy Protection Techniques

1. Face Anonymization: One of the most popular methods aims at covering or eliminating identity details such as:
   - Blurring, Pixelating or masking the face
   - Replacement of real faces with synthetic or cartoonized face
2. Skeletal and Pose-Based Representation: Instead of RGB Images and videos, systems can rely on:
   - Human skeletons
   - Pose keypoints

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

II. CONCLUSION AND FUTURE WORK:

Human monitoring and detection systems are significantly improved with the advancement of visual computing methods, learning algorithms, and neural network-based techniques, which are able to perform automated tasks such as attendance tracking, behavior analysis, monitoring of employees, smoking detection, and occupancy analysis in academic, office, and public settings. Moving from the traditional human-based monitoring systems to deep learning and other combined models has enhanced the accuracy and real-time capabilities of the systems.

To move forward, there is a need to focus on multimodal and cross-domain models, including Edge-AI models that are fast to run in real-time, as well as privacy-friendly and explainable models. Adding behavior analysis, increasing dataset diversity, and creating a benchmark will help to increase efficiency and useability.

REFERENCES

[1] Mol, P. A., & Mathew, T. E. (2025, June). Human Action Recognition–Analysis of Deep Learning Methods. In *2025 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI)* (pp. 1-7). IEEE.

[2] Feichtenhofer, C. (2020). X3d: Expanding architectures for efficient video recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 203-213).

[3] Arisandi, D., Elveny, M., & Rahayu, R. (2021, June). Human Detection and Identification for Home Monitoring System. In *Journal of Physics: Conference Series* (Vol. 1898, No. 1, p. 012026). IOP Publishing.

[4] Khan, A., Somani, I., & Kumari, E. (2025, March). Cam-Based Human Activity Recognition for Remote Patient Monitoring and Abnormality Alert System. In *2025 6th International Conference on Recent Advances in Information Technology (RAIT)* (pp. 1-7). IEEE.

[5] Kamal, U., Ahmed, S., Toha, T. R., Islam, N., & Alim Al Islam, A. B. M. (2020). Intelligent human counting through environmental sensing in closed indoor settings. *Mobile Networks and Applications*, *25*(2), 474-490.

[6] Ansari, M. A., & Singh, D. K. (2021). Human detection techniques for real time surveillance: a comprehensive survey. *Multimedia tools and applications*, *80*(6), 8759-8808.

[7] Sulman, N., Sanocki, T., Goldgof, D., & Kasturi, R. (2008, December). How effective is human video surveillance performance?. In *2008 19th International conference on pattern recognition* (pp. 1-3). IEEE.

[8] Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human factors*, *50*(3), 433-441.

[9] Paul, M. (2013). *Human detection in surveillance videos and its applications.* EURASIP Journal on Advances in Signal Processing (review).

[10] Ula, M., Pratama, A., Asbar, Y., Fuadi, W., Fajri, R., & Hardi, R. (2021, April). A new model of the student attendance monitoring system using rfid technology. In *Journal of Physics: Conference Series* (Vol. 1807, No. 1, p. 012026). IOP Publishing.

[11] Al Hajji, K., Cenadi, A., & Ahmad, F. (2022, June). Machine learning based human face recognition for attendance system. In *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-7). IEEE.

[12] Sakthikumar, B., Ramlaingam, S., Saghana, T., Sahaya Joseph Idin, G., Shakunthala, S., & Shivadharshan, S. (2025, April). Smart AI Based Attendance Monitoring System Using YOLOV8. In *2025 3rd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)* (pp. 1-6). IEEE.

[13] Deepa, T., Krishna, K. T., Chandra, P. S., & Kumar, K. S. (2023, May). An Intelligent Computer Vision based Application for Student Monitoring in Classroom. In *2023 4th International Conference for Emerging Technology (INCET)* (pp. 1-7). IEEE.

[14] Raj, A., Srivastav, H., Shukla, S., & Gupta, N. (2024, March). Facial Recognition-Based Student Attendance System. In *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 1-5). IEEE.

[15] Rajesh, T. R., Kuchipudi, L. N., & Dammalapati, A. (2024, July). Attendance System Based on Facial Recognition Using OpenCV. In *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)* (pp. 1529-1535). IEEE.

[16] Umar, U. S., Rana, M. E., & Hameed, V. A. (2023, December). Predicting Student Performance Through Classroom Concentration: A Review on Machine Learning and Computer Vision Approaches. In *2023 IEEE 21st Student Conference on Research and Development (SCOReD)* (pp. 408-414). IEEE.

[17] Issac, T., & Terrance, J. S. (2024, June). Investigation of Student Engagement Monitoring System using Machine Learning. In *2024 Second International Conference on Inventive Computing and Informatics (ICICI)* (pp. 47-52). IEEE.

[18] Dudhmal, R., Khatik, I., Kadam, S., Choudhary, S., Zurange, S., & Borate, V. (2025, March). Monitoring Students in Online Learning Environments Using Deep Learning Approach. In *2025 3rd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)* (pp. 202-206). IEEE.

[19] Al-Sakib, A., Islam, F., Haque, R., Islam, M. B., Siddiqua, A., & Rahman, M. M. (2024, February). Classroom activity classification with deep learning. In *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1-7). IEEE.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

[20] Ami, F. I. A., Yumang, A. N., & Quinto, E. J. M. (2025, March). Student Posture Extraction and Attentiveness Classification Using Deeplabcut and Support Vector Machine. In *2025 17th International Conference on Computer and Automation Engineering (ICCAE)* (pp. 59-62). IEEE.

[21] habib Albohamood, A., Alqattan, M. S., & Vizcarra, C. P. (2025, April). Real-time Student Engagement Monitoring in Classroom Environments using Machine Learning and Computer Vision. In *2025 4th International Conference on Computing and Information Technology (ICCIT)* (pp. 420-424). IEEE.

[22] Zheng, R., Jiang, F., & Shen, R. (2020, May). Intelligent student behavior analysis system for real classrooms. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 9244-9248). IEEE.

[23] Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109.

[24] Akash, S., Dinesh, V., Selvan, S. M., Daniel, J. A., & Srikanth, R. (2023, May). Monitoring and Analysis of Students' Live Behaviour using Machine Learning. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 98-103). IEEE.

[25] Hamidi, N. H. H. M., Abidin, N. A. Z., Aminuddin, R., Sheng, C. C., Samah, K. A. F. A., & Nasir, S. D. N. M. (2024, September). A Real-Time System for Monitoring Student Drowsiness in the Classroom Using the Deep Learning Model YOLOv8. In *2024 5th International Conference on Artificial Intelligence and Data Sciences (AiDAS)* (pp. 105-110). IEEE.

[26] Cui, K., Huang, M., Lv, W., Liu, S., Zhou, W., & You, Q. (2024). Research on Intelligent Recognition Algorithm of College Students' Classroom behavior based on improved Single Shot Multibox Detector. *IEEE Access*.

[27] Santoso, Y. D., & Tjahyanto, A. (2024, December). Modification of YOLOv9 with Bi-Level Routing Attention for Student Behavior Monitoring in Crowded Classrooms. In *2024 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 537-542). IEEE.

[28] Monti, L., Tse, R., Tang, S. K., Mirri, S., Delnevo, G., Maniezzo, V., & Salomoni, P. (2022). Edge-based transfer learning for classroom occupancy detection in a smart campus context. *Sensors*, *22*(10), 3692

[29] Sha, O. (2022, September). Student detection and people counting system based on apriori algorithm. In *2022 International Conference on Education, Network and Information Technology (ICENIT)* (pp. 138-142). IEEE.

[30] Challa, K., Sharma, A., Darwish, H., AlHmoud, I. W., Islam, A. K., Graves, C., ... & Gokaraju, B. (2024, March). Optimizing HVAC efficiency via deep neural networks for real-time classroom occupancy. In *SoutheastCon 2024* (pp. 735-738). IEEE.

[31] Harshavardhan, G., Singh, N. T., & Chanu, A. D. (2023, December). Real-time employee monitoring system using CCTV cameras. In *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1-7). IEEE.

[32] Nazarov, D., & Nazarov, A. (2023, June). Intelligent service for monitoring the activities of an employee of an organization. In *2023 IEEE 25th Conference on Business Informatics (CBI)* (pp. 1-5). IEEE.

[33] Kumar, A., & Prabavathy, A. K. (2025, May). Hybrid Multimodal AI Framework for Real-Time Employee Activity Recognition and Workplace Motion Analytics. In *2025 Fourth International Conference on Smart Technologies, Communication and Robotics (STCR)* (pp. 1-8). IEEE.

[34] Gopal, P. G., Subhash, N. S., & Anil, R. A. (2022, August). Analysis of Employee Surveillance System using Deep Learning Models. In *2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA* (pp. 1-6). IEEE.

[35] Shriraam, K., & Deepa, N. (2023, May). An innovative application for employee attendance using near field communication to reduce the time complexity using IP and geo tracking comparing with biometrics. In *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)* (pp. 1-6). IEEE.

[36] Memane, R., Jadhav, P., Patil, J., Mathapati, S., & Pawar, A. (2022, August). Attendance monitoring system using fingerprint authentication. In *2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA* (pp. 1-6). IEEE.

[37] Pawar, A., Hiwanj, R., Koparde, P., Chikmurge, D., & Barve, S. (2023, September). Automated Employee Attendance Monitoring Using Liveness Face Recognition and Geofencing in Real Time. In *2023 IEEE Fifth International Conference on Advances in Electronics, Computers and Communications (ICAECC)* (pp. 1-7). IEEE.

[38] Harshavardhan, S., Upender, P., Ganesh, N., & Giftsy, S. (2025, January). SmokerBeacon: Hybrid Real-Time Smoking Detection Using Object Detection and Gas Sensor. In *2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (pp. 1-6). IEEE.

[39] Li, M., & Liu, F. (2024, November). A Novel Finetuned YOLOv8 Transfer Learning Model for Smoking Behavior Detection. In *2024 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML)* (pp. 1944-1949). IEEE.

[40] Jin, Z., Zai, W., Huang, J., & He, Y. (2023, April). Monitoring of smoking behavior in construction sites of the power system. In *2023 Panda Forum on Power and Energy (PandaFPE)* (pp. 1471-1475). IEEE.

[41] Sathiskumar, N. R., Sivagnanasekar, G., Vijayarajan, S., & Shameem Mohamed, A. R. (2024, December). Smoke Sentry: Employing YOLO-based Deep Learning for Sophisticated Real-Time Smoking (Cigarette smoking) Detection in Surveillance Networks. In *2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS)* (pp. 892-898). IEEE.

[42] Patwal, A., Diwakar, M., Tripathi, V., & Singh, P. (2022, June). Crowd counting model using convolutional neural network. In *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)* (pp. 156-160). IEEE.

[44] Blangiardi, F., Beltrán, R., Zhao, Z., Reinhard, S., Röβler, M., Langer, J., ... & Kuhn, H. (2024, November). Exploring Data Fusion for AI-Based People Counting in a Static Public Transport Vehicle Using Multiple IR-UWB Transceivers. In *2024 Sensor Data Fusion: Trends, Solutions, Applications (SDF)* (pp. 1-8). IEEE.

[45] Yang, Y., Hu, P., Shen, J., Cheng, H., An, Z., & Liu, X. (2024). Privacy-preserving human activity sensing: A survey. *High-Confidence Computing*, *4*(1), 100204.

[46] Liu, Y., Liu, S., Zhu, X., Yang, H., Li, J., Guo, J., ... & Liu, J. (2025). Privacy-preserving video anomaly detection: A survey. *IEEE Transactions on Neural Networks and Learning Systems*.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)