iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82542

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

CCTV Network-Based Crowd Management and Work Monitoring System Using AI/ML

Mrs. P. Neelima[1], M. Himasri[2], K. Harshitha[3], CH.Bhanu Tejaswini[4]
[1]*Assistant Professor, Department of Computer Science & Engineering, Bapatla Women's Engineering College, Bapatla, Andhra Pradesh,522101*
[2,3,4]*B. Tech, Department of CSE(AIML), Bapatla Women's Engineering College, Bapatla, Andhra Pradesh,522101*

Abstract: The current project aims to use the possibilities of existing Closed-Circuit Television (CCTV) networks to manage crowds and monitor the workplace using the latest Artificial Intelligence (AI) and Machine Learning (ML) methods. As the population in the urban areas continues to increase and the security issues continue to rise, there is a great need to ensure that the already existing CCTV networks are efficiently used. Conventional surveillance systems are highly dependent on manual surveillance which besides being inefficient is also subject to human error. The implementation of AI and ML in CCTV surveillance systems is an innovative solution to improve the management of the crowds and surveillance in the workplace. It is a system that uses real-time video analysis to identify anomalies, monitor suspicious events, and streamline the productivity of the workforce. The proposed solution will guarantee proactive surveillance, lower the workload of human operators, build enhanced security and efficiency due to the implementation of AI in the sphere of object recognition, behavior analysis, and predictive analytics. The use of the deep learning method, including YOLO, facilitates the automatic recognition of abnormal behavior, unauthorized entry, and safety violations in the workplace. Moreover, predictive algorithms assist in the optimization of the flow of the crowd, which makes the public space safer, and the working environment more efficient.
Keywords: CCTV Networks, Crowd Management, Workplace Monitoring, YOLO Object Detection, Anomaly Detection, Deep Learning, behavior analysis, Public Safety, AI/ML.

I. INTRODUCTION

The evolution of technology has led to significant advancements in public safety and operational monitoring. The aspect of Artificial Intelligence (AI) and Machine Learning (ML) integration with the current CCTV systems has become one of the pillars of developing safer and more efficient environments. It takes advantage of the proliferation of surveillance cameras to turn the conventional monitoring process into a proactive, intelligent, and automated process using this system. Although CCTV networks are crucial, they are usually limited to real-time surveillance and analysis of complicated scenes. Human operators may have difficulty with processing streams of information that are continuously on them and thus they may take long to respond and monitor critical events. Such issues become particularly conspicuous when it comes to the areas of crowd management, where timely detecting the overcrowding could help avoid accidents, or monitoring the workplace, where violations of the safety will have to be identified quickly. The capability to use the current CCTV infrastructure is one of the best characteristics of this system as it dramatically decreases the costs and the barriers to its implementation. At live video feeds, AI and ML algorithms can react and identify important events using installed cameras. Not only does it reduce the requirements of extra hardware, but this method also guarantees scalability and flexibility in different environments like in the open space, workstations, and high-security zones. The suggested system has two major features, that is, crowd management and work monitoring. In the case of crowd management, complex object detection algorithms such as YOLO are used to detect and estimate the number of individuals and categorize the density as normal or overcrowded. As a work monitoring system, the system will guarantee that safety measures and productivity are observed allowing organizations to maintain a high level of operational efficiency via automated AI surveillance.

The detection of cyberbullying and harassment on social media is based on machine learning, and computer vision is used to monitor the physical space. Equally, the methodology is based on automated detection of human [3]behavior and patterns of movements and safety breaches on CCTV channels, especially in open areas like a transport center, business quarters, and workplaces. The classification modules are designed to run within a monitored environment to identify all kinds of [1]behavioral trends such as crowds congestion, violations of restricted areas, and non-adherence to workplace safety. This is done by using deep learning classifiers, which are convolutional architectures. These classifiers operate within a supervised system based on annotated video data collections in order to identify the existence of abnormal activityin an observed scene.

The given approach allows the system to effectively detect cases of overcrowding, unauthorized access, and productivity, which comes to the continuous attempts to enhance the safety of the population and efficiency at the workplace.

II. LITERATURE REVIEW

Surveillance-based anomaly detection has been an active research area, and the integration of AI into CCTV systems has accelerated significantly over the past decade.

Smith et al. [1] have presented an in-depth discussion of the incorporation of intelligent video surveillance systems as a means to improve the security in urban areas, by stating that the AI-based decision support systems are able to process large volumes of CCTV data in real-time, thereby identifying and classifying potential security threats faster and to a greater degree than ever before.

According to Patel et al. [2], CCTV networks are a vital tool that helps to observe the crowd dynamics, overcrowding, and security risks. They have written about smart surveillance cameras that can study the behavior of a crowd, and some of the solutions are crowd density monitoring, queue monitoring, crowd flow, and incident detection related to the crowd, which provides basic standards of the density-based classification.

The author Johnson [3] discusses the application of the ML classifiers like Support Vector Machines (SVM) and Random Forest to identify abnormalities in surveillance videos, which have high accuracy on detecting anomalies and can be considered to be scalable to large data sets, but fail to work effectively in complicated settings without adequate training data. This restriction is the reason why deep learning methods are employed in the suggested system.

Suggested by Gupta [4], YOLO models are used in real-time with the view of measuring crowd density and monitoring. The research exhibits rapid and dynamic object recognition in different environments. Weaknesses are error reduction in dark conditions and high overlap cases with objects, which the proposed system deals with the technique of preprocessing and data enhancement.

The review of Wang et al. [5] covers machine learning methods of video surveillance anomaly detection, such as SVMs, Random Forests, clustering, and autoencoders. It is tested on real-world data, and its implications of the study to security monitoring and detecting new behavioral patterns within surveillance networks are discussed.

In identifying suspicious behaviors, Brown [6] uses CNNs and RNNs to carry out behavioral analysis, which supports behavior analysis in real-time with high accuracy. The research would need huge labelled datasets and can be biased in its model, and thus inclined towards the issue of varied annotated training materials which is provided by the proposed system in the form of public surveillance datasets.

Sharma [7] introduces AI-oriented systems to track the workings of industries, improve efficiency, and gain safety compliance. An important issue identified is the problem of privacy and excessive reliance on automation. We have solved these issues in the proposed system by use of anonymization of data and human-in-the-loop alert controls.

Li et al. [8] address the complex nature of urban security and surveillance in the fast-paced urbanization. They also emphasize the fact that CCTV networks are inherently vulnerable to several challenges including limited coverage, data saturation, and ineffective monitoring behaviors and suggest the introduction of AI/ML to
detect and analyze suspicious activities in real-time, which will be the primary rationale behind this project.

Zhan et al. [9] present a comprehensive review of the methods of crowd analysis, including the density estimation approaches, counting, and motion analysis approaches in various settings. Their taxonomy of the crowd analysis tasks acts as a direct influence in the design of the crowd management module in the proposed system. Zhang et al. [10] also add cross-scene counting of crowds through deep convolutional neural networks, which allow the strong generalization of viewpoints in a camera and the environment under various conditions.

Wojke et al. [11] introduce Deep SORT that builds on the SORT tracking algorithm by incorporating a deep appearance-based approach to minimize any identity switch in the process of occlusion. They are based on the person tracking part of the proposed system since their Kalman filter motion prediction and cosine-distance re-identification allow person tracking over congested frames with persistent assignment of its ID.

Chalapathy and Chawla [12] perform an extensive overview of the deep learning algorithms used in detecting anomalies, including autoencoders, recurrent networks, and generative adversarial models that have been used in different fields such as video surveillance. They analyzed the use of LSTM-based sequence classifier to identify temporal deviation in the movement patterns, which directly underlies the design of the behavioral anomaly detection module in the proposed system. The proposal of
Nath et al. [13] is a YOLOv3-based deep learning system that detects personal protection equipment on construction sites in the real-time which has better accuracy in identifying hard hats and safety vests. The result of their findings supports the use of single-stage detectors to PPE monitoring and directly report to the workplace safety compliance module of the proposed system.

The paradigm of single-stage multi-scale detection presented by Redmon and Farhadi [14] under the name YOLOv3 forms the basis of the further extended versions of the paradigm known as YOLO. Their experiment on estimating bounding boxes with three scales based on Darknet backbone illustrates the speed-accuracy trade-off that can make YOLO architectures the default model in real time surveillance systems including the one suggested here.

Long Short-Term Memory networks are proposed by Hochreiter and Schmidhuber [15], which addresses the vanishing gradient issue of the traditional RNNs, and is capable of learning long-range temporal relationships. The basis of their work is the behavioral anomaly detection module of the suggested system that is based on the LSTM sequence classification to detect unusual movement patterns under the study of multi-second surveillance windows.

He et al. [16] propose ResNet, The proposed system uses pre-trained ResNet-50 weights through transfer learning on the PPE compliance classifier which enables high-performance generalization of the workplace safety dataset using limited labeled training subjects.

Hasan et al. [17] introduce a convolutional autoencoder model to unsupervised anomaly detection in surveillance video, where normal compression actions are learned, and reconstructions with large errors are considered to be anomalous. The proposed system makes use of the supervised LSTM methodology, which provides a complementary view to the methodology offered by the authors, emphasizing the value of various training data to ensure strong detection across types of scenes.

III. SYSTEM ANALYSIS

The approach used in this work is structured to take the form of a pipeline which includes collection of data, model development, system integration and real-time deployment. The general design philosophy will focus on flexibility to the currently deployed CCTV infrastructure, thus making sure that the solution proposed is capable of being implemented without the need to enhance hardware or require network upgrades at a considerable cost. The system is designed to be implemented in a limited computational platform with an acceptably low level of detection latency in realistic security environments. The process of development is separated into four phases in a row.

The initial step will be critical evaluation of the current methods of surveillance to find out what they lack and ensure the functionality of the target system. The second step is data curation and preprocessing by creating the annotated video corpus required to train supervised models. The third step involves model choice, training, and optimization, during which deep learning models are evaluated by task-specific performance criteria. The fourth and the last stage involves the system integration, the end-to-end testing in the actual deployment environments and the validation of the performance relative to the baseline procedures.

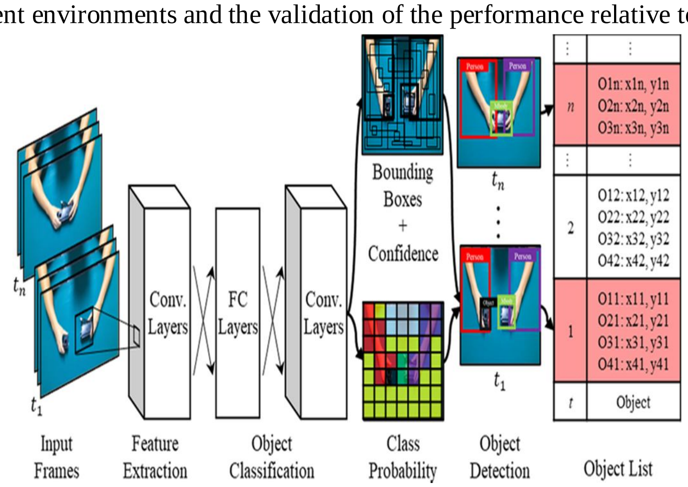

Figure 1: Model architecture

A. Existing System

In the current workplace and crowd monitoring systems, strategies of anomaly detection are based on rule-based thresholds, as well as the review of footage by a person. Such measures enable operators to detect blatant violations but are constrained by the amount of human attention and the processing lag. Nonetheless, with the size of the contemporary city monitoring systems, numerous surveillance functions are beyond human capacity and it presents lapses in real-time reaction. The state-of-the-art currently uses a number of computer vision algorithms: classic background subtraction methods to detect motion, Haar cascade and basic frame differencing methods to detect people and activities respectively.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Although these methods are computationally inexpensive, they have high false alarms in changing lighting conditions, and have no ability to classify [1]behavioral context. The current systems also do not combine the effects of the crowd density estimation with the monitoring of the workplace safety compliance, and standalone tools are needed. With these efforts, the existing systems are still lacking the accuracy to detect such intricate [1]behavioral changes as overcrowding patterns, unauthorized access events, and PPE non-compliance. The absence of a unified, scalable AI-powered system that can manage both crowd management and workplace surveillance at the same time is the main gap that this project will fill.

B. Proposed System

The suggested methodology is aimed at offering an all-inclusive approach to the detection of threats in CCTV images of the population and the workplace, such as breaches of crowd density, unauthorized entry, and workplace violation of safety. In order to achieve this objective, diverse deep learning models are used, such as object detectors based on [4]YOLO, temporal [15](LSTM-based) models, and [1]behavioral [6](CNN-based) classifiers.

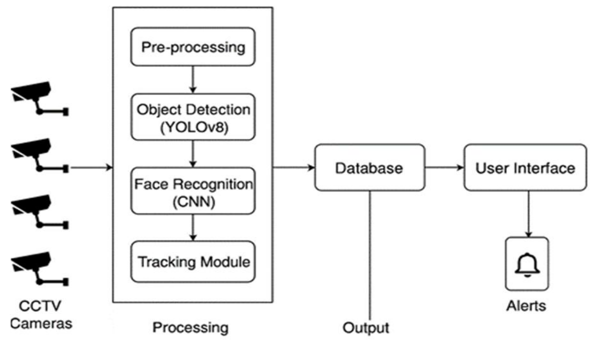

Figure 2: Proposed system architecture

The main components and characteristics of the proposed system will be: Training The system will be trained on annotated video datasets of both normal and anomalous surveillance environments with a variety of crowd densities and workplace conditions to provide a solid model performance.

The trained models will review real time live CCTV feeds and automatically identify violations, abnormal [1]behavior, overcrowding, unauthorized access and safety non-compliance and will trigger real time alerts to security personnel. Features of the application Python-based framework with the integration of OpenCV to process videos and TensorFlow and Keras to run deep learning models, including [13] YOLO and [17]LSTM.

• The backend architecture can handle multiple concurrent camera feeds and thus person detection and object tracking along with crowd density and workplace safety monitoring can be run concurrently. The system architecture entails a Sensor Layer, Video Processing Layer and AI Analysis Layer, an Alert Generation Layer, a Dashboard Layer, and a Sensor Layer.

• There is a Streamlit/Django dashboard that offers the real-time monitoring, visualization of heatmaps and crowd graphs, and access to analytics features reports to administrators. Event logging and storage of detected violations along with the maintenance of historical records of the surveillance are performed with the help of a PostgreSQL database.

• The notification module is a system that combines Telegram and SMS to provide instant notifications about the critical incidents so that the response to the incident could be timely and the situation could be controlled better.

C. Methodology

The methodology incorporates a multi-stage deep learning pipeline that processes live CCTV feeds by detecting, tracking, examining [3]behavior, and verifying safety compliance. It is implemented on top of Python 3.9, including OpenCV, TensorFlow 2.12, and Keras and can run on both edge hardware and GPU servers, as well as not require the upgrade of infrastructure.

1) Preprocessing: Preprocessing assists the raw CCTV recordings to be preprocessed before being fed to downstream analysis. All videos streams are partitioned into frames and reduced to 224x224 pixels and pixel values bounded between 0 and 1. Background subtraction and Gaussian blurring detect noise and isolate foreground objects whereas histogram equalization corrects uneven light. Per camera zone, regions of Interest (ROI) are used to limit processing to only what is important and minimize overhead.

2) Feature Extraction: Feature extraction has been used to encode discriminative pattern of preprocessed frames. The CSPDarknet53 backbone of [13]YOLOv5 detects spatial features and produces bounding box positions and class labels of each detected object. Through the Deep SORT tracker, motion features are obtained; centroid coordinates (x, y), velocity vectors, and trajectory angles. Temporal features are extracted by a two layer [15]LSTM encoding of 30 frame sequences, and the density of the crowd is estimated by summing person counts in each zone.
3) Model Selection: Model selection emphasizes on detecting, inference speed and hardware compatibility. [13]YOLOv5 is the backbone of object detection with a high map of about 45 FPS and person and PPE categories. ResNet-50 takes care of PPE compliance classification using ImageNet weights as a form of transfer learning. Two-layer [15]LSTM is used to record temporal abnormalities over 30 frames windows and [11]DeepSORT is able to offer effective multi-object tracking with deep appearances description coupled with Kalman filter prediction values.
4) Model Training: The data is stratified into training (70%), validation (15%) and test (15) sets. Initialization uses pre-trained ImageNet and COCO weights and then fine-tuning is performed using the CCTV-specific dataset. All models are trained by the Adam optimizer (lr=0.001) using cosine annealing and trained on the IOU loss, ResNet-50 on binary cross-entropy, and the LSTM on categorical cross-entropy. Early stopping (patience=10) helps to avoid overfitting, and training is done on the NVIDIA GPU using TensorFlow 2.12 and Keras.
5) Evaluation Metrics: The System quantitative measures are used to check the performance of the system. Quality of classification is gauged by Accuracy, Precision, Recall and F1-Score, where the latter is especially relevant to the skewed classes in anomaly detection. Mean Average Precision (map at 0.5 IoU) is a metric that is used to evaluate object detection, whereas Frames Per Second (FPS) is a metric that is used to test real-time feasibility. False Alarm Rate (FAR) and Confusion Matrix are checked to achieve operational feasibility of all system modules

IV. DATASET

1) Crowd Density: This classification organizes video frames in accordance with the amount of people in the area; it goes as far as wasting it to mini-crowd and excessive. With these annotations, the real-time estimation of the crowd and the generation of alerts using thresholds are possible.
2) Anomalous Behavior: This category will contain footage containing unusual [3]behavior as loitering, unauthorized access, violent movement, discarded items, and spontaneous surges of crowds. It aids the [2]behavioral analysis module in the determination of safety critical events.
3) Workplace Safety Compliance: This category consists of annotated data of PPE identification (hard hat, safety vests, and masks), area restriction offenses, and productivity controls. It identifies the obedient and the disobedient [3]behavior of workers.
4) Normal Activity: Frames of normal pedestrian movement and normal working activity are used as the reference point in anomaly detection models.
5) Dataset Split and Augmentation: The dataset will be split into Training (70%), Validation (15%), and Testing (15) sets. Flipping, brightness, and rotation are also data augmentation methods used to enhance robustness in different environmental conditions.
6) Motion Pattern Analysis: Directional flow, abrupt motion changes, running, or pattern of crowd dispersal Annotated sequences depicting directional flow, abrupt motion changes, running behavior, or pattern of crowd dispersal. These samples help to be used in the training of temporal models of [6]behavioral anomalies.

V. IMPLEMENTATION AND RESULTS

The implementation of the system was based on Python 3.9 and OpenCV 4.7, TensorFlow 2.12 and Keras in the form of the deep learning pipeline. The [14] YOLO v5 model has been chosen to conduct real-time object detection because it has a good trade-off of both inference speed and detecting accuracy. This model was optimized using the annotated surveillance data presented in the previous section, and it was trained in 100 epochs on an NVIDIA GPU with Adam as the optimization algorithm and learning rate of 0.001.

A. Secure Vision About Page Overview

Secure Vision is introduced as a real-time AI-powered surveillance platform built on [14] YOLOv8 deep learning. The About page highlights its two core detection modules: crowd control and weapon detection. It emphasizes real-time processing with support for both live camera feeds and pre-recorded video.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

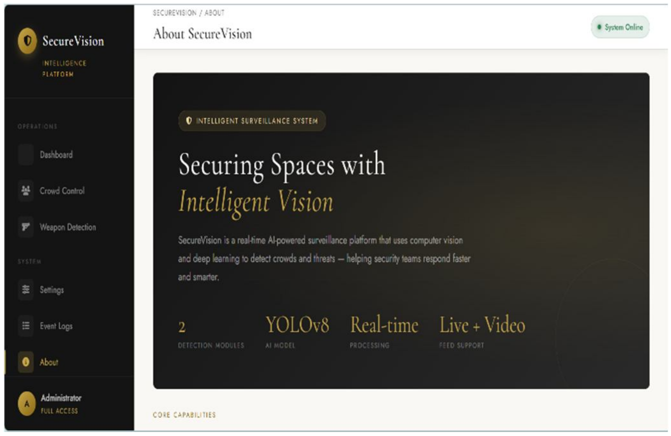

Fig 5.1 Secure Vision

B. System Configuration Alerts & Audio Settings

The Settings page allows administrators to configure audio alerts for each detection module separately. Crowd and weapon alerts are mapped to distinct MP3 files loaded from the static/audio/ directory. Both alert sounds are shown as active, ensuring audible notifications fire on detection events.

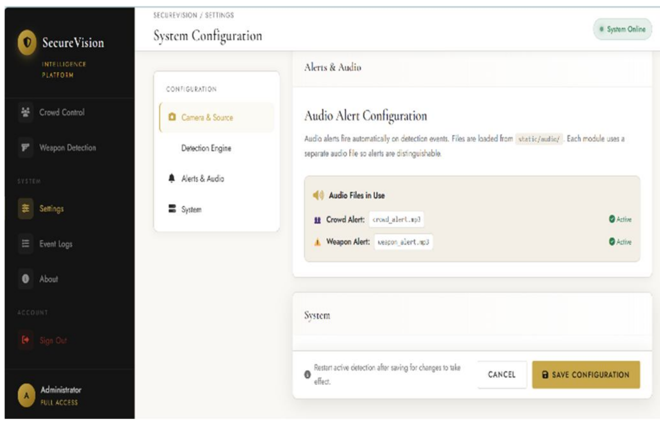

Fig 5.2 System Configuration Alerts & Audio Settings

C. Live Webcam Feed with Crowd Detection

The dashboard displays a live annotated webcam stream with a bounding box drawn around a detected person, triggering a "Crowd Alert" overlay. The Alert Feed confirms the system is idle between detections, while the Video Library panel shows one uploaded file available for analysis. The annotated feed mode visually highlights detected subjects in real time.

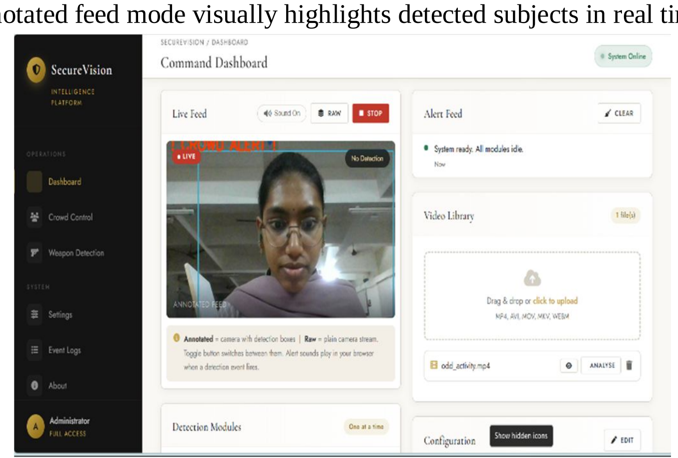

Fig 5.3 Live Webcam Feed with Crowd Detection

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

D. Video File Analysis Mode

The dashboard shows Secure Vision analyzing a pre-recorded video file (ODD_ACTIVITY.mp4) with crowd detection actively running. The live feed panel displays the video preview alongside real-time crowd alerts in the Alert Feed panel. Users can toggle between annotated and raw views during playback.

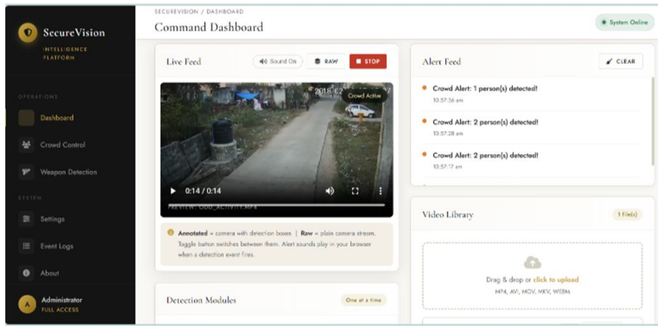

Fig 5.4 Video File Analysis Mode

E. Live Weapon Detection on Webcam Feed

The live annotated webcam feed shows the system actively identifying a knife held in front of the camera, with the "Weapon Active" status badge displayed on the stream. The Alert Feed records two consecutive weapon detection events — knife at 90% and 64% confidence — within seconds of each other. A browser-level popup notification also appears at the top right, confirming real-time alerting when a weapon detection event fires.

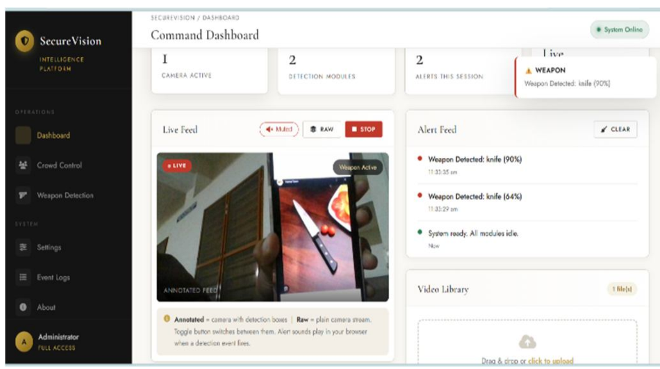

Fig 5.5 Live Weapon Detection on Webcam Feed

F. Event Logs Detection & System Events Timeline

The Event Logs page displays a timestamped history of 211 system and detection events in chronological order. Entries are categorized by type — Alert or Detection — with color-coded badges for quick identification. This log serves as an audit trail, validating the system's continuous monitoring and alerting capability throughout a session.

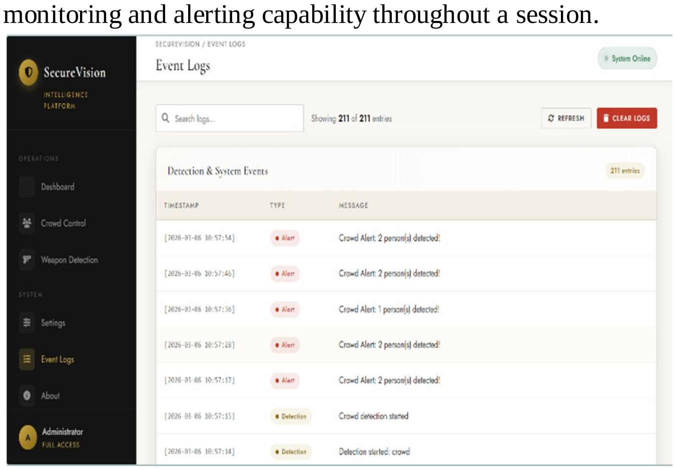

Fig 5.6 Event Logs Detection & System Events Timeline

1) Crowd Detection Results: On a high-level and busy urban environment, the model has detected and labeled 14 people in a single frame and the scene is categorized as a Mini Crowd according to the density thresholds that are predefined. The background-coloured bounding boxes were drawn around each identified individual and this established the correct detection under a high-traffic setting. The crowd data graph monitored the number of crowds, number of violations, event of restricted entry and abnormal crowd [7]behavior as time moved and this gave the security people a real time perspective of the situation.
2) Monitored Work: The PPE compliance detection was found to be 91 percent accurate on the test data on workplace safety. The live alert mechanism was able to send a notification through Telegram in an average period of 1.2 seconds after the anomaly had been detected. Live heatmaps, crowd count timelines and violation logs were shown on an interactive Streamlit dashboard, which enabled operators to watch multiple camera feeds at once. All detection events were stored in the centralized PostgreSQL database by the timestamps which made them valuable in the past to analyze the history and identify the trend and use this to plan security.
3) Performance Evaluation: The system was tested and evaluated on a rigorous test and proved to be reliable, adaptive and responsive to different security scenarios. The crowd detection model was based on [14] YOLO and had a mean Average Precision (map) of 0.83 and the IoU threshold of 0.5 on the test set. The [8]behavioral anomaly classifier was found to have a total accuracy of 87.4, and precision of 84.2 and a recall of 89.1. False alarm rate was kept at less than 8% which proves to be practical in real life application. The results of Table I provide the comparison of the offered system performance with the current baseline methods.

VI. CONCLUSION

In address the increased demand of smart and scalable city surveillance, this project suggests a new AI-based CCTV surveillance system that will merge crowd control with occupational safety surveillance into one real-time system. The system upgrades passive surveillance infrastructure to an active safety management system using neural networks and neural network-based [12]behavioral classifiers, such as [13] YOLO object detectors, [17] LSTM time-based analysis, and CNN-based behavioral classifiers.

The results of the experiment provide a clear understanding of the fact that the deep learning-based technique is significantly better than the conventional rule-based and classical machine learning techniques in terms of all the evaluation metrics, such as accuracy, precision, recall, F-measure, and false alarm rate. The capability of the system to work with the existing CCTV equipment, without the need to install more cameras is an important cost savings factor to the practical implementation of the system in an urban setting. Moreover, the study aims not only at counting crowds but also at detecting [2]behavioral abnormalities and ensuring compliance with safety requirements at workplaces, which takes into consideration the variety of safety necessities of contemporary urban and industrial settings.

Real-time alerting, interactive dashboards, and historical analytics enable the security personnel and facility managers to have complete situational awareness. The AI-based surveillance despite the obstacles such as computational needs in high-density views, variability in performance when it is in the dark, and privacy concerns related to continuous surveillance cameras are significant improvements in the field of intelligent, efficient, and proactive surveillance. The privacy controls that were integrated comprise data anonymization, role-based access control, and encrypted data storage to ensure that they are deployed responsibly within the regulatory cover.

VII. FUTURE WORK

Even though at present the system shows good performance, it still has a lot of opportunities to improve and be expanded. Future evolution directions consist of the implementation of edge computing with shorter inferences times to allow the system to run on embedded hardware with no need to use centralized servers, which would significantly enhance scalability to large-scale smart city deployments. Privacy-preserving model training Federated learning is another direction of significant importance as it allows models to be trained using distributed camera networks without sending raw video data to central servers.

Multi-camera cross-scene analysis with enduring person re-identification across non-overlapping camera views would greatly increase the potential of the system tracking persons of interest across a large physical area. Optimization of models by implementing neural architecture search and model compression methods such as quantization and pruning would make real-time execution on resource-constrained edge devices. The addition of smart city features such as traffic control systems and emergency response platforms would facilitate coherent automated reaction to identified safety incidents, allowing even more significant latency in human reaction.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

REFERENCES

[1] Smith, J. AI-Powered Surveillance Systems for Urban Safety. IEEE Transactions on Smart Cities, 2019.
[2] Patel et al. Crowd Management in Public Spaces: Challenges and Solutions, 2020.
[3] Johnson, E. Machine Learning for Anomaly Detection in Surveillance. ACM Computing Surveys, 2020.
[4] Gupta, R. Crowd Management Using YOLO-Based Object Detection. International Journal of Computer Vision, 2021.
[5] Wang et al. Machine Learning Approaches for Anomaly Detection in Video Surveillance, 2021.
[6] Brown, M. Deep Learning Models for Behavior Analysis in Surveillance. Pattern Recognition Letters, 2023.
[7] Sharma, P. AI-Driven Work Monitoring Systems in Industrial Environments. IEEE Transactions on Industrial Informatics, 2023.
[8] Li et al. Urban Security and Surveillance: Challenges and Opportunities, 2017.
[9] Zhan, B. et al. Crowd Analysis: A Survey. Machine Vision and Applications, vol. 19, pp. 345–357, 2008.
[10] Zhang, C. et al. Cross-Scene Crowd Counting via Deep Convolutional Neural Networks. IEEE CVPR, Boston, 2015.
[11] Wojke, N.; Bewley, A.; Paulus, D. Simple Online and Realtime Tracking with a Deep Association Metric. In Proc. IEEE Int. Conf. Image Process. (ICIP), Beijing, China, 2017; pp. 3645–3649.
[12] Chalapathy, R.; Chawla, S. Deep Learning for Anomaly Detection: A Survey. arXiv preprint arXiv:1901.03407, Jan. 2019.
[13] Nath, N.D.; Behzadan, A.H.; Paal, S.G. Deep Learning for Site Safety: Real-Time Detection of Personal Protective Equipment. Autom. Constr., vol. 112, p. 103085, Apr. 2020.
[14] Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767, Apr. 2018.
[15] Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. Neural Comput., vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[16] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016; pp. 770–778.
[17] Hasan, M.; Choi, J.; Neumann, J.; Roy-Chowdhury, A.K.; Davis, L.S. Learning Temporal Regularity in Video Sequences. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016; pp. 733–742.

AUTHORS

Mrs. P. Neelima, M. Tech, Assistant Professor, Dept of computer Science & Engineering, BWEC, Andhra Pradesh, India.
Palepuneelima31@gmail.com

M. Himasri, [B. Tech], Student, Dept of Computer Science & Engineering (AI & ML), BWEC, Andhra Pradesh, India.
himasrimallisetty2004@gmail.com

K. Harshitha, [B. Tech], Student, Dept of Computer Science & Engineering (AI & ML), BWEC, Andhra Pradesh, India.
harshithakota88@gmail.com

CH. Bhanu Tejaswini, [B. Tech], Student, Dept of Computer Science & Engineering (AI & ML), BWEC, Andhra Pradesh, India.
B hanuchejarla7@gmail.com