Journal of Ubiquitous Computing and Communication Technologies (ISSN: 2582-337X)
www.irojournals.com/jucct/

A Hybrid Intelligent Crowd Monitoring System for Real-Time Public Safety

Manikandaraja G.[1], Jaisanth M.[2], Balaji M.[3]

[1]Assistant Professor, School of Mechanical Engineering, SRM Institute of Science and Technology, Tiruchirappalli, India

[2,3]Student, Department of Bachelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Tiruchirappalli, India

E-mail: [1]bgmanikandaraja@gmail.com, [2]jaisanthmani@gmail.com, [3]balajimanikandan902@gmail.com

Abstract

Crowd monitoring has become an essential component of public safety management through continuous monitoring and detection of potential risks in congested areas. In this research, a hybrid intelligent crowd monitoring system comprising YOLOv8-based person detection, crowd density estimation, crowd flow analysis, abnormal crowd behavior detection, and alert generation is introduced. The design and implementation of the proposed hybrid intelligent crowd monitoring system have been accomplished utilizing the DMADV methodology in order to set up a defined design and implementation process. First of all, video frames from the surveillance cameras undergo pre-processing in order to improve their quality and then analyzed using YOLOv8 algorithm for person detection. Finally, the detected persons will be used for the estimation of crowd density and flow analysis as well as abnormal crowd behavior detection in real time. If the predetermined safety thresholds are exceeded, the automatic alert is generated.

Keywords: YOLOv8, Define-Measure-Analyze-Design-Verify (DMADV), Crowd Density Estimation, Person Detection, Crowd Flow Analysis, Intelligent Surveillance.

DOI: https://doi.org/10.36548/jucct.2026.3.004
Extended version of article received from **International Conference on Smart Innovations in Engineering & Technology (ICSIET 2026)**
Received: 22.05.2026, received in revised form:25.06.2026, accepted:13.07.2026, published: 25.07.2026

1. Introduction

Rapid growth in urban population and increasing frequency of large-scale public gatherings have increased the requirement for crowd management and safety surveillance systems. Urban areas like transport stations, shopping complexes, educational facilities, stadiums, and other public event venues encounter the presence of dense crowds that can give rise to congestion, restricted movement, security risks, and emergencies. Traditional surveillance systems heavily rely on continuous monitoring by humans of video feeds from multiple cameras, which may be inefficient due to delayed response time and human fatigue. The recent advancements in the field of computer vision and deep learning have made it possible to create intelligent surveillance systems that can detect, track, and analyze people in video feeds automatically.

To address these challenges faced by traditional surveillance techniques, this research presents the development of the Hybrid Intelligent Crowd Monitoring System for Real-Time Public Safety. In the proposed system, several functionalities of surveillance are combined into a single platform including person detection, crowd density calculation, entrance-exit detection, crowd flow, and abnormal behavior detection. A YOLOv8 object detection algorithm is used to detect persons in real time, and the crowd analytics module calculates the occupancy rate and the movement pattern of the crowd. Moreover, an alert generation module and monitoring dashboard have been developed which provides the security staff with information related to overcrowding and unsafe behavior.

The proposed Hybrid Intelligent Crowd Monitoring System is an integrates the YOLOv8-based person detection, crowd density estimation, crowd flow analysis, abnormal activity detection, and automated alert generation into comprehensive surveillance system for real-time monitoring of public safety issues. The surveillance system is designed using the Six Sigma DMADV methodology to guarantee systematic processes of design, implementation, and validation of the surveillance system. The performance of the system is experimentally validated using a customized crowd image dataset which includes publicly available online images and surveillance images from various public places.

2. Literature Review

Real-time object detection act as a fundamental component of modern crowd monitoring systems by providing precise localization and detection of objects in surveillance environments. One-stage detectors provided an efficient balance between precision of detection and computational efficiency and were thus well-suited for large scale surveillance applications [1]. Further developments involved better techniques for feature extraction and multi-scale predictions, which provided more accurate detection of smaller, occluded, and tightly packed objects [2]. Region-based and feature pyramid detectors, on the other hand, provided better localization at different scales; however, they had increased computational cost, which restricted them from being used in real-time crowd monitoring applications [5], [6].

The crowd density estimation has received considerable attention through the use of Convolutional Neural Networks (CNNs) to handle difficult cases like severe occlusions and scale variations. Multi-column CNNs enhanced the accuracy of density estimation by making feature extraction from various receptive fields. Furthermore, several reviews have demonstrated the efficiency of deep learning methods in crowd counting under difficult environments [3], [4]. However, the main concern of these methods is density estimation.

Analysis of crowd behavior and anomaly detection has also received an attention in terms of the identification of abnormal behaviors within crowded settings. Deep learning techniques have been utilized for modeling crowd behavior and identification of unusual incidents using surveillance videos, which helps in early detection of any security threats [7], [8], [9]. Furthermore, techniques for crowd counting and profiling have helped in understanding crowd characteristics for surveillance applications [10]. Nevertheless, most researches that have been performed till date on the above-discussed issues have focused only on one aspect such as detection, density estimation, or anomaly detection. This research proposes a hybrid intelligent crowd monitoring system that integrates various techniques within a single framework.

3. Methodology

3.1 System Architecture

The design of the hybrid intelligent crowd monitoring system provides a real-time surveillance and public safety monitoring of crowded environments. It consists of video preprocessing, person detection using YOLOv8, crowd density analysis, crowd flow analysis, abnormal behavior detection, and automated alert generation as modules. This design ensures continuous monitoring of the public space and facilitates the identification of possible threats.

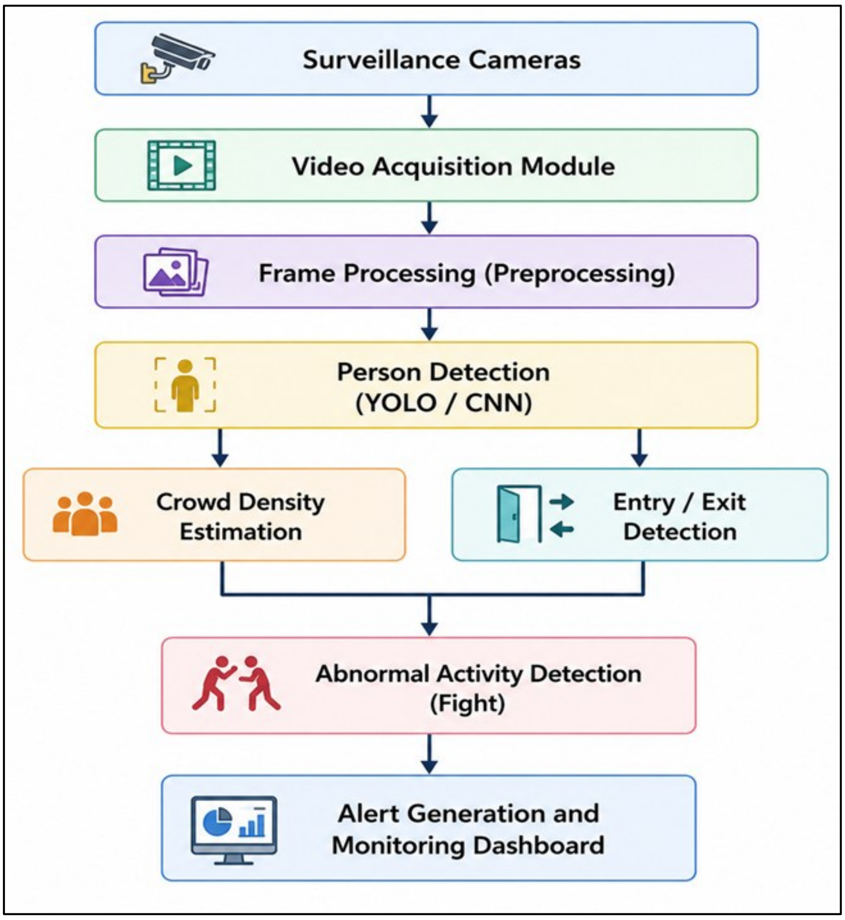


Figure 1. Architecture of the hybrid intelligent crowd monitoring system

As shown in Figure 1, the frames of the surveillance videos are captured and preprocessed to improve the image quality and prepare the data for further processing. The processed images are further supplied to the YOLOv8 model to detect and localize individuals in the scene. The detected individuals are used for the crowd density estimation and crowd flow analysis, whereas the abnormal crowd behavior is monitored all the time to discover any unusual events. Depending on the obtained results, the alert generation component sends alerts in case of any detected crowd density threshold or abnormal behavior.

3.2 Architecture of the YOLOv8 Detection Model

The proposed hybrid intelligence-based crowd monitoring system uses the YOLOv8 object detection architecture as the foundation for real-time detection of persons in crowds. The overall architecture of the adopted YOLOv8 is illustrated in Figure 2. It comprises four main elements: Backbone, Neck, Detection Head, and Post-processing. The Backbone extracts spatial features in the form of a hierarchy using convolutional and C2f blocks, while the SPPF Spatial Pyramid Pooling - Fast (SPPF) block helps extract information in the form of receptive field. The Neck fuses features at multiple scales using top-down and bottom-up streams for better detection of persons in crowds at different scales. The Detection Head makes predictions in terms of bounding boxes, objectness score, and persons class label, while the post-processing stage uses confidence thresholds and Non-Maximum Suppression (NMS).

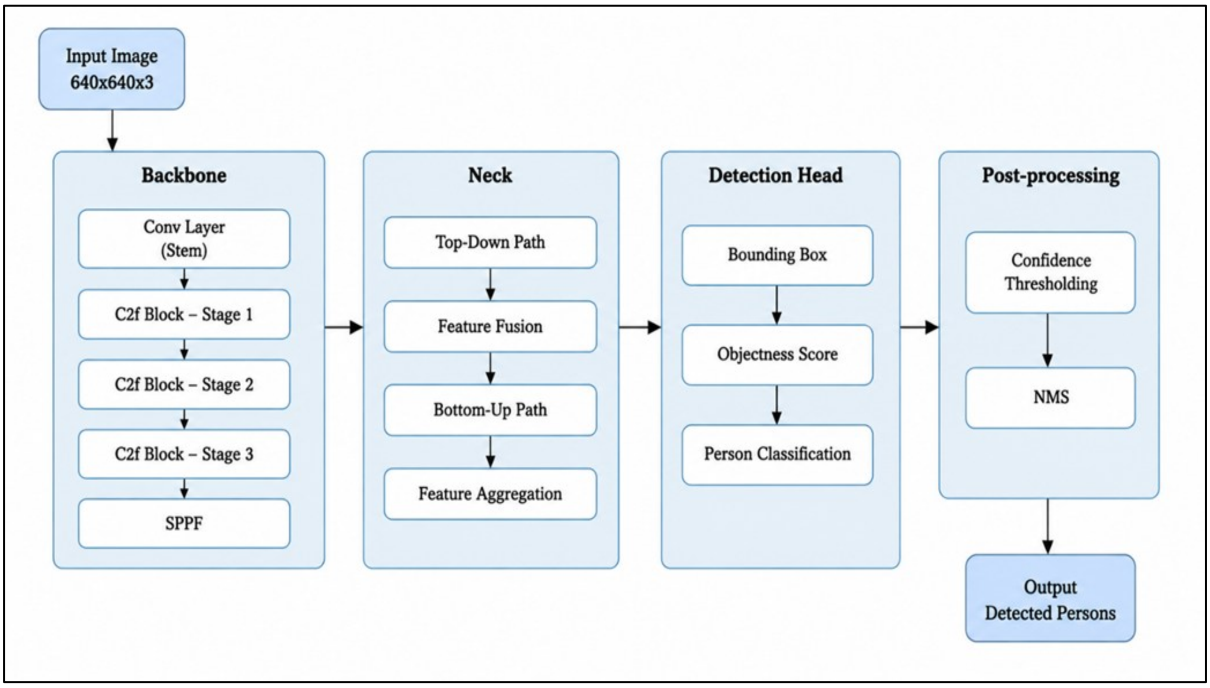


Figure 2. YOLOv8 architecture adopted for the proposed crowd monitoring system

YOLOv8 was selected because it provides a good trade-off in terms of accuracy of detections, speed of inference, and computational efficiency, which make it a proper solution for real-time crowd monitoring. This model uses an anchor-free approach to detection, which makes localization less complicated and increases the accuracy of detections, while the C2f modules help with feature extraction with fewer calculations. The feature aggregation approach used in the Neck helps in detecting people at different crowd densities and scales, as well as partially occluded objects. Moreover, the architecture allows real-time inference with regular GPU, which is needed for detecting crowded areas and abnormal activity in order to monitor public security.

The preprocessed video frame is denoted as I. The YOLOv8 algorithm predicts the bounding box and confidence values of the detected persons as,

$$P_d = \{(B_i, C_i)\}_{i=1}^{N} \tag{1}$$

where B_i denotes the bounding box of the i^{th} detected person, C_i represents the confidence score, and N is the total number of detected persons. NMS is then applied to eliminate redundant detections.

Algorithm 1: Overall crowd monitoring workflow

Input: Live Video Stream

Output: Crowd Statistics and Alerts

Begin

Step 1: Capture video frames

Step 2: Pre-process frames

Step 3: Detect persons using YOLOv8

Step 4: Count detected persons

Step 5: Step Track detected individuals

Step 6: Estimate crowd density

Step 7: Monitor entry and exit crossings

Step 8: Analyze crowd movement patterns

Step 9: Detect abnormal activities

Step 10: Generate alerts if thresholds exceeded

Step 11: Display results on dashboard

End

3.3 Dataset Description

The data was collected from public crowd scenes and consists of various places, crowd density, viewpoints, and lightings. In total, 1,450 crowd images were collected and examined manually for ensuring data quality and variation of scenes. The entire data was annotated for the person detection problem by using the bounding box according to the YOLO data annotation method. The collected data was separated into training and test parts for experiments based on the 80:20 ratio. Thus, 1,160 images were utilized for training the YOLOv8, and 290 images were left for testing the performance of the model. Figure 3 illustrates some images from the data collected for this research.

Figure 3. Sample images used for crowd monitoring

3.4 DMADV-based Design Approach

The Hybrid Intelligent Crowd Monitoring System was designed employing the DMADV model of 6 Sigma to create a structure in designing and validating the system. In the Define stage, the problems of traditional surveillance systems and specifications for intelligent crowd monitoring were defined. The Measure stage involved defining some critical performance indicators in the areas of detection accuracy, crowd analysis, and responsiveness. The Analyze stage focused on assessing the traditional surveillance models to define the issues of crowd density, occupancy, and abnormality detection.

Figure 4. DMADV-based development framework for the proposed crowd monitoring system

The design stage included the integration of person detection, crowd density estimation, tracking, crowd flow analysis, and alarm generation in an integrated way. Finally, verification was done on the performance of the system using crowds and video streams from surveillance cameras. The entire process followed for the development of this system based on DMADV is shown in Figure 4.

3.5 Video Acquisition and Pre-processing

Input images were collected from publicly available sources and real-world surveillance images taken in different public environments. It contains 1,450 crowd images that vary with respect to the density of crowds, lighting, viewpoint, and complexity of background. Before using the dataset for training, all images were converted into 640×640 pixels to meet the input requirements of the YOLOv8 model. The pixel values were normalized, and data augmentation such as horizontal flip, scaling, and brightness adjustments were used to enhance the model's generalization capacity under surveillance conditions.

Algorithm 2: Frame pre-processing

Input: Video Frame
Output: Enhanced Frame
Begin

Step 1: Resize frame
Step 2: Convert color format
Step 3: Apply noise filtering
Step 4: Adjust brightness and contrast
Step 5: Normalize image
Step 6: Return processed frame

End

3.6 Person Detection

The process of person detection is achieved using the YOLOv8 object detection model training, as shown in Figure 5. The input of each image is provided to the model where the persons are detected by predicting the bounding box and the associated confidence score. The predicted bounding boxes are further improved by applying the NMS algorithm to remove overlapping detections and keep the best possible results. The detected persons are then passed on to the crowd density estimation, crowd flow analysis, and abnormal activity detection modules. The entire process of person detection is shown in Figure 5.an ESP32 microcontroller to trigger alerts or safety measures. If the driver is alert, no response is generated (Fig 2).

Figure 5. Person detection using YOLOV8-based object detection

---

Algorithm 3: Person detection

Input: Pre-processed Frame

Output: Detected Persons

Begin

- **Step 1:** Load YOLOv8 model
- **Step 2:** Pass frame to model
- **Step 3:** Detect human objects
- **Step 4:** Generate bounding boxes
- **Step 5:** Assign confidence scores
- **Step 6:** Return detected persons

End

3.7 Crowd Density Estimation

Crowd density estimation is performed based on the number of people detected in the surveillance area. The total count of people detected is continually calculated and then matched to predefined occupancy threshold values to identify crowd density. The crowd is classified based on the level of occupancy as either normal, moderate, or critical. In case the crowd detected exceeds the predefined threshold, an overcrowding alarm is sent whenever the number of people detected is above the predetermined value.

The crowd density is estimated from the total number of detected persons within the monitored area.

$$D = \frac{N}{A} \tag{2}$$

where,

- $D =$ crowd density (persons/ m^2)
- $N =$ total detected persons
- $A =$ monitored area
- A higher density value indicates increased crowd congestion.

Algorithm 4: Crowd density estimation

Input: Number of Detected Persons
Output: Density Status
Begin

Step 1: Count detected persons
Step 2: Compare count with threshold
Step 3: If count > threshold
- Generate overcrowding alert

Step 4: Else
- Continue monitoring

Step 5: End If

End

3.8 Tracking, Crowd Flow Analysis, and Abnormal Activity Detection

Following person detection, unique IDs are assigned to the detected persons to ensure continuous tracking across multiple frames. Tracking will enable the determination of movement trajectories, hence avoiding double counting while conducting the occupancy analysis. A virtual boundary is then set in the surveillance area to keep track of the entry and exits depending on movement direction. The obtained trajectory information is used to study the movement pattern of crowds to identify areas that are congested or unusual movement behavior.

The track information is further analyzed to study motion properties and interaction among people. Crowd flow analysis is done based on the movement vectors from sequential frames in order to identify the direction as well as the occupation pattern of the people. Moreover, any kind of abnormal behaviors, like aggressive behavior or change in motion property, is detected by the analysis of motion pattern. Alerts for any of the behaviors exceeding a certain threshold are generated and sent to the monitoring dashboard.

Movement of the detected individual is calculated based on the displacement between two consecutive images.

$$V_i = \frac{\sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}}{\Delta t} \tag{3}$$

where,

- V_i = velocity of the i^{th} person
- (x_t, y_t) = current position
- (x_{t-1}, y_{t-1}) = previous position
- Δt = time interval between two frames
- This computation enables the estimation of crowd movement patterns.

Abnormal crowd behavior is detected in case of high values for the estimated crowd density or crowd movement.

$$A = \begin{cases} 1, & D > D_{th} \text{ or } V_i > V_{th} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where,

- $A = 1$ indicates abnormal activity
- D_{th} = density threshold
- V_{th} = movement threshold

Algorithm 5: Tracking, crowd flow analysis, and abnormal activity detection

Input: Detected Person Coordinates and Bounding Boxes

Output: Entry/Exit Statistics, Crowd Flow Information, and Alert Status

Begin

Step 1: Assign unique identifiers to detected individuals.

Step 2: Track individual positions across consecutive frames.

Step 3: Record entry and exit events using a virtual boundary.

Step 4: Generate movement trajectories for tracked individuals.

Step 5: Analyze movement vectors to estimate crowd flow patterns.

Step 6: Identify congestion regions and unusual movement behavior.

Step 7: Evaluate motion characteristics and interaction patterns.

Step 8: Detect abnormal activities based on predefined behavioral thresholds.

Step 9: Generate alerts when abnormal events or threshold violations are detected.

Step 10: Forward crowd statistics and alert information to the monitoring dashboard.

End

3.9 Alert Generation and Monitoring Dashboard

The monitoring dashboard functions as the primary interface for displaying crowd analytics and surveillance information through the use of the proposed system. Data collected using person detection, crowd density calculation, tracking, crowd flow calculation, and abnormal behavior detection blocks are all compiled and made available on the dashboard. The dashboard continuously gives information on the occupancy levels, number of people entering and leaving the place, crowd movements, and events detected.

The process of improving situational awareness has been achieved through the addition of an automated alerting system. Whenever certain thresholds concerning crowding, density, or abnormal behavior exceed set values, alerts are created by the system. Overcrowding instances, strange movement patterns, and abnormal behaviors are flagged immediately in the dashboard to facilitate appropriate action. Through the integration of real-time visualization and automation alerts, the dashboard has made crowd management more efficient.

Generation of the alert takes place when abnormal crowd behavior is detected.

$$\text{Alert} = \begin{cases} ON, & A = 1 \\ OFF, & A = 0 \end{cases} \quad (5)$$

Alerts are sent to the monitoring system to help the security officers intervene in time.

4. Results and Discussion

4.1 Experimental Setup

The proposed framework for hybrid intelligent crowd monitoring was developed using the programming language Python along with OpenCV and Ultralytics YOLOv8. Experimental evaluation of the proposed framework was done on a computer with an Intel Core i7 processor, NVIDIA GPU, and 16 GB RAM to provide support for real-time inference and surveillance processing. Input images for training were scaled to 640 × 640 pixels, and preprocessing such as normalization and data augmentation was carried out before training of the proposed model. A custom image dataset for crowds with a total of 1,450 images was collected from the internet and real-world crowd monitoring scenarios and used to perform evaluation of the proposed framework. The image dataset was divided into 80% training dataset (1,160 images) and 20% test dataset (290 images). System evaluation was performed by Accuracy, Precision, Recall, and F1-score metrics. Qualitative analysis of the proposed system was carried out on person detection, crowd density calculation, crowd flow analysis, abnormal behavior detection, and automatic alerts.

4.2 Person Detection Performance

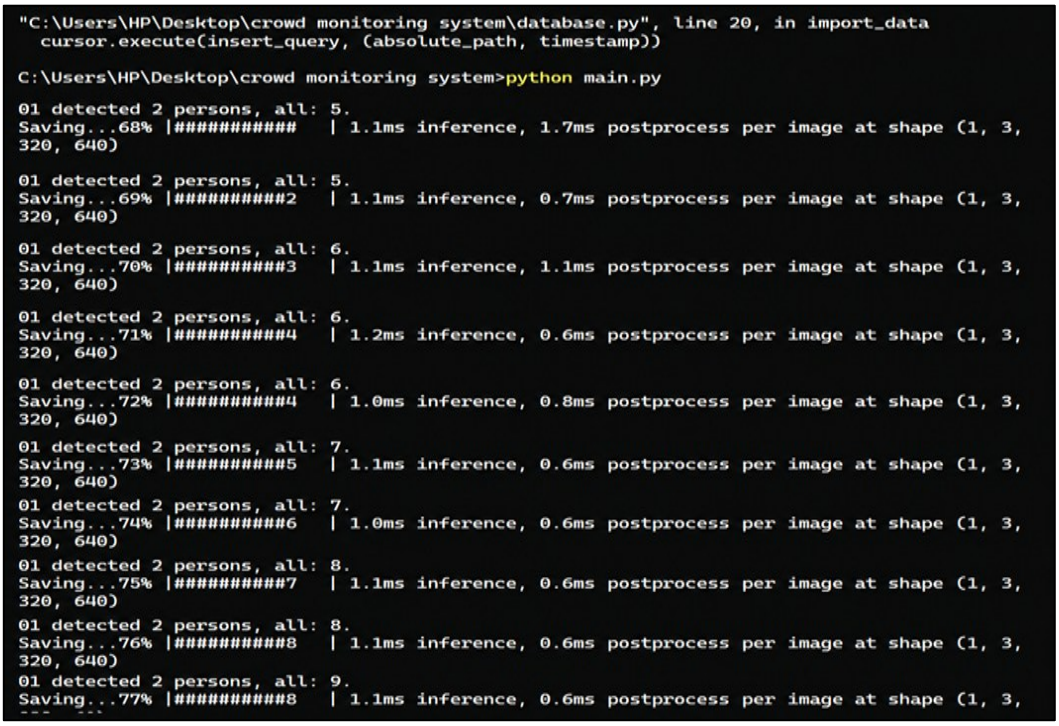
The person detection system was effective at detecting people in surveillance scenes with different levels of crowd density and motion. Localization of the subjects allowed for accurate estimation of the number of occupants as well as serving as the basis for the other processes in crowd analysis. The bounding boxes and confidence values helped in detecting people without duplication using non-maximum suppression.

The detection results demonstrated the ability of the system to ensure consistent performance within the actual surveillance environment, since there were continuous changes in the composition and movement patterns of the crowds. Accuracy in person detection is necessary in crowd monitoring systems, since mistakes in this module have a direct effect on density estimation, tracking, and behavior analysis modules. Figure 4 shows the outputs for person detection from the surveillance video streams.

4.3 Crowd Density Analysis

The crowd density measurement module efficiently performed crowd analysis using the density of occupants in the surveillance area by assessing the number of occupants. Using certain predetermined density standards, crowd behavior could be classified into varying levels of occupancy, making it possible to differentiate between the presence of normal, moderate, and heavy crowds.

The crowd density data generated helped in crowd management proactively as it made it possible to identify areas where the occupancy was high. The monitoring of crowd density would lead to alerts when the thresholds were crossed hence enabling immediate action and preventing any congestion-related accidents from taking place. Figure 6. shows the crowd density monitoring and detection of overcrowding.



```
"C:\Users\HP\Desktop\crowd monitoring system\database.py", line 20, in import_data
  cursor.execute(insert_query, (absolute_path, timestamp))

C:\Users\HP\Desktop\crowd monitoring system>python main.py

01 detected 2 persons, all: 5.
Saving...68% |###########   | 1.1ms inference, 1.7ms postprocess per image at shape (1, 3,
320, 640)

01 detected 2 persons, all: 5.
Saving...69% |##########2   | 1.1ms inference, 0.7ms postprocess per image at shape (1, 3,
320, 640)

01 detected 2 persons, all: 6.
Saving...70% |##########3   | 1.1ms inference, 1.1ms postprocess per image at shape (1, 3,
320, 640)

01 detected 2 persons, all: 6.
Saving...71% |##########4   | 1.2ms inference, 0.6ms postprocess per image at shape (1, 3,
320, 640)

01 detected 2 persons, all: 6.
Saving...72% |##########4   | 1.0ms inference, 0.8ms postprocess per image at shape (1, 3,
320, 640)

01 detected 2 persons, all: 7.
Saving...73% |##########5   | 1.1ms inference, 0.6ms postprocess per image at shape (1, 3,
320, 640)
01 detected 2 persons, all: 7.
Saving...74% |##########6   | 1.0ms inference, 0.6ms postprocess per image at shape (1, 3,
320, 640)
01 detected 2 persons, all: 8.
Saving...75% |##########7   | 1.1ms inference, 0.6ms postprocess per image at shape (1, 3,
320, 640)
01 detected 2 persons, all: 8.
Saving...76% |##########8   | 1.1ms inference, 0.6ms postprocess per image at shape (1, 3,
320, 640)
01 detected 2 persons, all: 9.
Saving...77% |##########8   | 1.1ms inference, 0.6ms postprocess per image at shape (1, 3,
```

Figure 6. YOLOv8-based person detection and crowd monitoring results

4.4 Tracking, Crowd Flow Analysis, and Abnormal Activity Detection Results

The tracking module successfully maintained the tracking history of detected persons between consecutive video frames. This made it possible to continuously track the crowd movements in the surveillance environment. The created trajectories were useful for occupancy tracking by capturing instances of people entering and exiting the scene, and gave information about the pattern of movements and the location of the crowd members.

The crowd flow analysis identified the trends of movements and variations in occupancies in various regions within the monitored area. With the help of extracted movements, it was possible to identify the regions that had higher concentrations of crowds along with certain flow directions. This information was helpful for crowd management. Moreover, the abnormal activity detection component constantly analyzed the motions and interactions in order to identify unusual activities. Alert generation was used to ensure prompt detection of potential threats and quick response of the security personnel. Figure 4 provides some examples of the results obtained through tracking, crowd flow analysis and abnormal activity detection.

4.5 Overall System Performance

The overall, result of the performance evaluation clearly demonstrates that the crowd monitoring system designed by us operates efficiently and steadily when performing various surveillance functions such as people detection, crowd density estimation, occupancy detection, tracking, and abnormal behavior detection. It should be noted that due to the integration of the architecture, the coordination between the analytical modules is possible.

Figure 7. Real-time monitoring dashboard and alert interface

The performance analysis of the monitoring modules is presented in Figure 7. the system exhibits consistency in behavior for all types of surveillance activities, which shows that the framework operates reliably irrespective of the crowd density and situation.

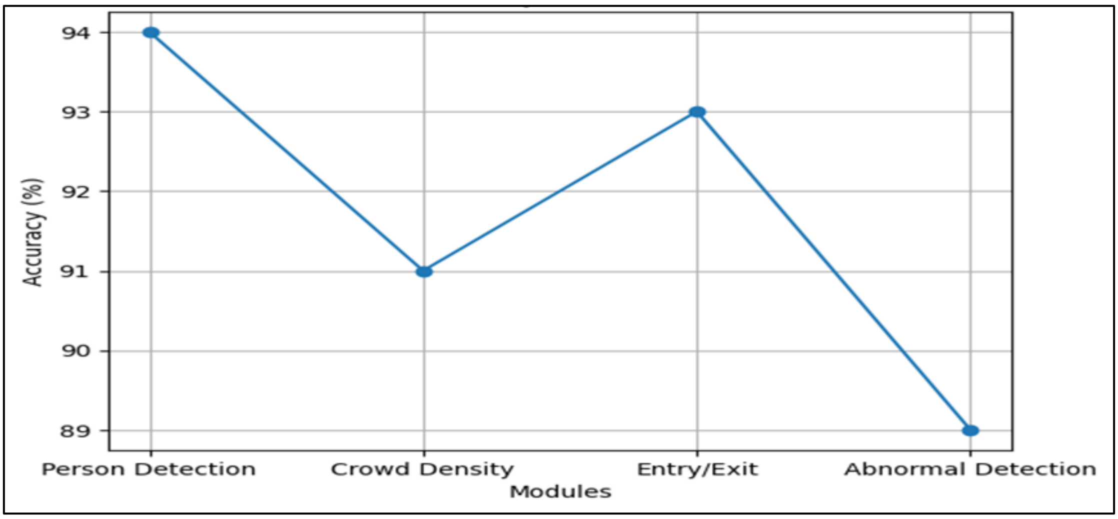


Figure 8. Comparative performance analysis of system modules

Figure 8 illustrates the trend of evaluation results for the proposed system. As indicated by the results in the figure, the stability of the framework in different monitoring processes is confirmed. The consistency of the performances of the various components analyzed in the figure suggests that the framework is capable of processing crowd-related.

Table 1. Performance evaluation of the proposed crowd monitoring system

Module	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Person Detection	94	92	93	92.5
Crowd Density Estimation	91	90	89	89.5
Entry/Exit Detection	93	91	92	91.5
Abnormal Activity Detection	89	87	88	87.5

The performance results quantitatively obtained from the important modules of the crowd monitoring system under consideration have been provided in Table 1. The findings shown in the table provide evidence for the results shown in Figure 7, thereby confirming the effectiveness of the developed framework. The findings suggest that the proposed method is capable of efficiently handling several crowd monitoring tasks in an intelligent surveillance system.

4.6 Discussion

The proposed hybrid intelligent crowd monitoring framework shows the integration of person detection, crowd density estimation, crowd flow analysis, abnormal behavior detection, and alert generation within the surveillance system. With the utilization of the YOLOv8 object detection technique, person localization is carried out efficiently in the crowded scenes. At the same time, density estimation and movement analysis provide an understanding of the dynamics of crowds. The development process of the framework through the utilization of DMADV methodology ensures reliable development processes and integration of the components. Compared to the existing surveillance systems where only one component is involved, this framework provides an opportunity to perform several operations like crowd monitoring and analysis and generate alerts. In addition, the application of the crowd dataset increases the practicality of the framework when it comes to the local deployment scenario.

5. Conclusion

The proposed hybrid intelligent crowd monitoring system provides a highly efficient framework for public safety through real-time surveillance and intelligent crowd analysis. The combination of YOLOv8-based person detection algorithm with crowd density estimation, crowd flow analysis, abnormal activity detection, and automatic alerts generation creates a unified framework for crowd monitoring purposes. The application of DMADV methodology ensures a well-structured design process while the use of a crowd image dataset collected from publicly available sources and real-life surveillance makes the framework more applicable for practical crowd monitoring cases. The proposed workflow allows for identifying possible threats in a timely manner which can be helpful for making fast decisions regarding crowd management. Future research will focus on integration of multi-camera surveillance, deployment of edge-AI for low latency inference, video understanding using transformer models, and predicting crowd behavior. Such changes will increase the scalability and robustness of the suggested framework, making it more applicable to the smart city surveillance systems.

References

[1] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, 2016: 779-788.

[2] Redmon, Joseph, and Ali Farhadi. "Yolov3: An Incremental Improvement." arXiv preprint 2018 arXiv:1804.02767.

[3] Zhang, Yingying, Desen Zhou, Siqin Chen, Shenghua Gao, and Yi Ma. "Single-Image Crowd Counting via Multi-Column Convolutional Neural Network." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016: 589-597.

[4] Sindagi, Vishwanath A., and Vishal M. Patel. "A Survey of Recent Advances in Cnn-based Single Image Crowd Counting and Density Estimation." Pattern Recognition Letters, ScienceDirect, 2018, vol 107: 3-16.

[5] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks." Advances in Neural Information Processing Systems 28 IEEE, 2015: 1-9

[6] Lin, Tsung-Yi, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. "Feature Pyramid Networks for Object Detection." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2017: 2117-2125.

[7] Zhou, Bolei, Xiaoou Tang, and Xiaogang Wang. "Learning Collective Crowd Behaviors with Dynamic Pedestrian-Agents." International Journal of Computer Vision, 2015, vol. 111, no. 1: 50-68.

[8] Sabokrou, Mohammad, Mahmood Fathy, Mojtaba Hoseini, and Reinhard Klette. "Real-Time Anomaly Detection and Localization in Crowded Scenes." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops,usingIEEE, 2015: 56-62.

[9] Tran, Tung Minh, Tu N. Vu, Nguyen D. Vo, Tam V. Nguyen, and Khang Nguyen. "Anomaly Analysis in Images and Videos: A Comprehensive Review." ACM Computing Surveys, 2022, vol. 55, no. 7: 1-37.

[10] Gao, M., Souri, A., Zaker, M., Zhai, W., Guo, X. and Li, Q., A Comprehensive Analysis for Crowd Counting Methodologies and Algorithms in Internet of Things. Cluster Computing, 2024, vol. 27, no.1: 859-873.