Smart Al-Driven Crowd Density Monitoring and Stampede Prevention System

* Madena Venu Gopala Rao
Undergraduate Student Dept of CSE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
Email: vtu22446veltech.edu.in
ph:9346772405

L.Kartheesan
Assistant Professor (Senior Grade) Dept of CSE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India Email:drkartheesanl@veltech.edu.in
ph:9884974007

Tholichuri Ravi Teja
Undergraduate Student Dept of CSE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
Email:raviteja.tholichuri@gmail.com

ph:9704144709

R. Vasanth Kumar Mehta et al. (eds.), *Proceedings of the International Conference on Intelligent Systems for a Sustainable Future (ISSF 2026)*, Atlantis Highlights in Intelligent Systems 16,
https://doi.org/10.2991/978-94-6239-693-7_54

Abstract. --Manual crowd monitoring at public gatherings, religious events, and transportation hubs often causes delays, inaccurate estimates of crowd density, and limited awareness of the situation, especially in busy environments where crowd behaviour changes quickly. This paper introduces a smart vision-based system that automatically estimates crowd density and gives early warnings to prevent stampedes using deep learning and real-time analytics. The system uses the YOLOv8-S model for person detection because it balances speed and precision well. To train the model, we used two different datasets: the public Shanghai Tech crowd dataset and our custom Crowd Safe dataset, which reflects various crowd densities, lighting conditions, and camera angles. Preprocessing modifies input frames to 640×640 pixels, while techniques like brightness adjustment, blurring, and perspective changes help the model work well across different scenarios. Detected individuals are analysed through a spatial clustering and heatmap-based density estimation module, which classifies crowd levels as low, moderate, or high. A risk assessment engine then combines temporal data with LSTM-based trend analysis to predict where overcrowding might occur. The system sends alert notifications to authorities and automatically activates control mechanisms like gate regulation or voice announcements. Experimental results show that the system accurately detects crowds and estimates density with more than 93% accuracy in different environmental conditions, making it suitable for live surveillance feeds. Current limitations include reduced performance in very low-light situations and cases of partial occlusion. Future research will aim to include thermal imaging and sensors to enhance robustness and improve early anomaly detection. This study shows that combining deep learning detection and temporal forecasting offers a practical and efficient way to manage crowds and prevent stampedes.

Keywords: : crowd monitoring, deep learning, density estimation, YOLOv8, LSTM, real-time analytics, surveillance, prevention, forecasting

1 INTRODUCTION

Handling public safety in crowded events has been increasingly difficult over the past few years. From religious congregations to protest rallies, sports events, and transportation terminals, dense public crowds are an everyday consequence of modern civilization. The Internet of Things has broadened the dimensions to upgrade and automate monitoring and crowd management systems. Smart surveillance systems, leveraging AI-powered analytics, can constantly scan video streams for patterns, anomalies, or outliers at speeds far greater than human operators can achieve. Among these, deep learning models-convolutional neural networks, or CNNs-have shown better performance in challenging visual recognition tasks by directly learning robust characteristics from the raw picture data. The family of object detection architectures known as YOLO has seen outstanding performance in real-time applications,

It provides a great balance between low latency and excellent precision. The latest variation of YOLOv8, with architectural improvements that increase small-object recognition, make training more stable, and reduce computational overhead, is perfectly suited for deployment on low-power embedded systems or edge surveillance cameras. Even with these developments, real-world crowd monitoring still poses a number of particular challenges. Lighting conditions, partial occlusions, camera viewpoint distortion, and rapid crowd movement can seriously impact the reliability of detection. Besides, since object detection provides only static estimates of crowd size, it cannot ensure effective management of crowd safety on its own. Temporal analysis should be included in a comprehensive monitoring system in order to identify patterns, predict traffic accumulation, and issue timely notifications before hazardous density thresholds are reached. Our study addresses these challenges by developing an integrated. It offers an excellent balance between high precision and low latency. With architectural enhancements that improve small-object detection, stabilize training, and lower computational overhead, the most recent iteration of YOLOv8 is ideal for deployment on edge surveillance cameras or low-power embedded systems. Despite such progress, a number of unique challenges persist in real-world crowd monitoring. These include rapid crowd movement, partial occlusions, distortion due to the camera view, and lighting conditions, all of which have significant impacts on the reliability of detecting objects. Secondly, object detection alone cannot be considered effective in crowd safety management, since it provides only static crowd size estimates. The system for comprehensive monitoring should include temporal analysis in order to identify trends, predict the building-up of traffic, and thus allow the issuance of warnings with due timing before dangerous density levels are reached. Our work addresses these challenges by proposing an integrated. Smart AI-Driven Crowd Density Monitoring and Stampede Prevention System. The proposed framework couples YOLOv8-based people detection with temporal analysis by Long Short-Term Memory to predict crowd density variations over time. By spatially clustering the detected people, the approach generates real-time heatmaps that visualize density distribution across any monitored zones. It continuously assesses the crowd dynamics and, when critical levels of density are detected, it may automatically trigger warning notifications or crowd-control actions such as voice alerts, gate regulation, or directional signage. The contribution of this work to the field of intelligent crowd management is manifold. First, it presents a scalable pipeline for training deep learning models on heterogeneous datasets representing diverse environmental conditions and camera viewpoints. Second, it combines real-time detection with predictive analytics, transforming raw surveillance data into actionable safety intelligence. Third, it demonstrates how AI-driven automation can supplement traditional human monitoring and mitigate reaction time while improving situational awareness in case of high-density events. Finally, the paper outlines the limitations of the system, including sensitivity to extreme occlusion and illumination conditions, and discusses a number of future extensions, such as multimodal sensor fusion and illumination-invariant feature extraction, that can improve robustness. The proposed system provides an example of how modern AI, combined with domain-specific safety knowledge, can realize a proactive solution to counter the menace of crowd disasters and improve the efficiency of public event management in smart cities with connected infrastructure

2 LITERATURE REVIEW

Zhang et al. [1] presented a deep-learning-based framework for crowd counting using CSRNet, which involved the utilization of dilated convolution layers in order to enhance feature extraction for dense populations. Their model achieved highly accurate results on the Shanghai Tech dataset and thus showed very strong generalization across different crowd densities, laying the foundation for real-time and accurate density estimation in surveillance systems.

Li et al. [2] combined a convolutional LSTM for temporal forecasting and YOLOv7 for crowd localization. This hybrid model overcame the drawbacks of static crowd detection by combining analysis of temporal sequences to improve the accuracy of congestion prediction and offer an early warning capability in crowded settings such as stadiums and metro stations.

Rahman et al. [3] designed an edge-based AI monitoring system for Internet of Things camera networks using a lightweight MobileNetV3 backbone that showed its promise for low-power, distributed deployment in real-world applications of surveillance by employing quantization and pruning, which reduced computing cost by roughly 50% with accuracy higher than 90%

Kumar et al. [4] proposed a hybrid approach that embodied YOLOv5-based person detection and captured the dynamics of crowd motion through optical flow analysis. The model would further detect abrupt directional shifts and other abnormal motion patterns, reaching 92.4% event recognition accuracy. This multimodal technique enhanced situational awareness by combining spatial and temporal information, an approach valuable for real-time anomaly detection in dense gatherings.

Wang et al. [5] They suggested the Crowd Sense framework that reduced frame-to-frame noise by combining Kalman Filter-based temporal smoothing with density map estimate. The approach was useful in terms of temporal stability during continuous crowd tracking, and its evaluation in urban settings reported a low mean absolute error.

Tan et al. [6] improved the YOLOv8-S model by embedding CA modules that are used for the recognition of small objects under crowd occlusion and complex lighting. The model achieved a 4.8% increase in precision while maintaining real-time speed, showing that attention mechanisms are important under high-density monitoring conditions.

Hossain et al. [7] designed a predictive congestion alert system that integrates CNN-based feature extraction with Recurrent Neural Networks for real-time crowd flow forecasting. Their model effectively predicted crowd buildup and high-risk zones using short temporal windows, thus demonstrating the benefits of combining spatial detection with sequential prediction for proactive risk management.

Liu et al. [8] solved the problem of wide-area monitoring using a holography transformation-based multi-camera fusion approach. By stitching together overlapping feeds of cameras in a unified spatial map, their system achieved improved localization accuracy and scalability, extending its applicability to large venues and open public areas.

Chen et al. [9] proposed a real-time stampede prediction framework that integrated YOLOv8 for human detection with an SFM for behavioral simulation. The system predicted the beginning of panic by analyzing interpersonal distances and velocity vectors, achieving over 95% accuracy in high-density simulations, providing a pragmatic basis for pre-emptive crowd control.

Zhou et al. [10] developed a scheme for crowd detection in low light, integrating YOLOv7-tiny and Vision Transformer (ViT) fusion, and managed to retain the same detection accuracy down to as low as 15 lux illumination. It showed the feasibility of AI-based surveillance for nighttime outdoor events where visibility is a major limitation.

Rahul et al. [11] propose an integrated crowd safety management platform using YOLOv8 for detection and LSTM for density trend forecasting. The proposed system provides more than 94% accuracy with sub-second alert latency, thus confirming the use of deep learning and temporal modeling to prevent stampedes in real time for intelligent public safety management.

Cumulative results from these studies depict the evolution of AI-based systems for crowd analysis from static density estimation to predictive and real-time monitoring frameworks. Most existing approaches either focus on individual accuracy in detection or behavior forecasting. Our proposed Smart AI-Driven Crowd Density Monitoring and Stampede Prevention System combines both YOLOv8 real-time detection and LSTM-based temporal prediction to enable accurate and proactive crowd management in dynamic environments with changing conditions.

3 PROPOSED METHODOLOGY

The proposed methodology will integrate computer vision, deep learning, and predictive analytics for automatic monitoring, analysis, and real-time forecasting of crowd density to prevent stampedes proactively. The system will be comprised of various interconnected modules, which are designed to be scalable and accurate while being deployed in complicated public environments. The overall workflow consists of six major stages: dataset preparation, preprocessing and augmentation, model training and optimization, real-time detection and density estimation, temporal prediction and alert generation, and system integration for user interaction.

A. Dataset Preparation

The system creates a strong training foundation from two complementary datasets: the Shanghai Tech Crowd Counting Dataset and the specially created Crowd Safe dataset. While the Shanghai Tech dataset offers various real-world crowd scenarios annotated with density maps, the Crowd Safe dataset was created to incorporate more event-specific photos-such as religious gatherings, concerts, and transit terminals. All photos were standardized into a.jpg format, while the annotations were transformed into YOLO-compatible text files with bounding box coordinates and class labels. The data was split into subsets for testing (10%), validation (20%), and training (70%) in order to build and evaluate the model..

B. Data Preprocessing and Augmentation

These steps ensure that the model is resilient against changes in viewpoint and environment. The images were first normalized and downsized uniformly to 640 by 640 pixels, then underwent several augmentations such as random rotation, horizontal flipping, modifications in brightness and contrast, and Gaussian blurring. This helps the model tackle real-world problems in surveillance feeds, like occlusion, perspective distortion, and variable lighting conditions. To make sure that the YOLOv8 network interprets the input consistently during training, a special YAML configuration file was created to map class labels (individual, group, and crowd).

C. Model Architecture and Training

The architecture of YOLOv8-S was adopted because it offers a good balance between real-time performance and high accuracy, which is needed to handle real-time applications on edge devices. Transfer learning was used with the pre-trained base weights yolov8s.pt to fine-tune the combined dataset. Training was conducted for multiple epochs with early stopping, adaptive learning rate scheduling, and mixed precision training to optimize convergence speed.

The training consisted of two phases:

1. Base Training Phase: Training on the Shanghai Tech dataset for general human detection and crowd distribution pattern learning.

2. Refinement Phase: Fine-tuning on the Crowd Safe dataset in order for the model to adapt to localized conditions like camera angles, event types, and lighting variability.

Mean Average Precision at 0.5, precision, recall, and F1-score were recorded as performance metrics using Tensor Board visualization during training.

D. Real-Time Detection and Density Estimation

Once trained, the YOLOv8-S detects persons in every frame and assigns a bounding box with corresponding confidence scores. The coordinates from the detected persons are mapped to a spatial density heatmap using Gaussian kernel estimation. In this map, crowd concentration is visualized by three threshold-based categories:

- Low Density (< 30%)
- Moderate Density (30–70%)
- High Density (> 70%)

The density estimation module updates in real time, and security operators can monitor changing conditions through visual overlays and live analytics dashboards.

E. Temporal Prediction and Alert Generation

To go beyond static detection, an LSTM-based temporal prediction module analyzes the crowd density history. The LSTM takes as input the sequence of densities extracted from previous frames and predicts the short-term crowd growth trend. If the predicted density level exceeds predefined safety thresholds, the system automatically triggers an alert mechanism.

Alerts can be generated through various channels:

- On-screen visual warnings on the dashboard.
- Voice announcements through automated systems to guide the movement of crowds.
- Signal activation for gate control or diversion in highly concentrated areas.

This predictive capability converts the system from a passive monitoring tool to an active early-warning mechanism for stampede prevention.

F. System Integration and User Interface

The entire pipeline is hosted inside a Flask-based web application, allowing seamless interaction between the end user and the AI modules. Users can either stream live feeds using the camera or upload any pre-recorded footage for processing. Its interface offers real-time outputs:

• Detection overlays: bounding boxes, heatmaps.

• Density level indication/color-coded warnings.

• Predictive alerts (textual and audible).

Integration with IoT control systems at the backend now enables automated crowd management actions like controlled entries or directional lighting. The modular architecture means that multi-camera fusion or thermal imaging can be integrated with ease in future versions of this system.

G. Summary

This approach incorporates real-time detection and spatial density estimation along with temporal prediction for proactive crowd management. Based on the detection via YOLOv8 and forecasting via LSTM, the system now provides instant situational awareness and short-term prediction critical to preventing overcrowding and ensuring public safety in dynamic large events. This will automate the recognition, classification, and measurement of bolts, nuts, and washers through appropriate warning techniques. The system is divided into distinct but interlinked modules and hence can easily be scaled up, interpreted, and adapted to real-world manufacturing conditions. In essence, the total workflow consists of five main stages: dataset preparation, data annotation and preprocessing, model training and optimization, measurement, and recommendation generation.

MATHEMATICAL FORMULATIONS

The proposed Smart AI-Driven Crowd Density Monitoring and Stampede Prevention System leverages deep learning, computer vision, and spatiotemporal analytics to estimate crowd density, detect abnormal congestion, and predict potential stampede scenarios in real time. The mathematical foundations governing crowd detection, density estimation, and risk assessment are presented below.

H. 1) Crowd Detection Formulation

The YOLO-based person detection model divides each input video frame into a grid of $S \times S$cells. Each cell predicts bounding boxes and their associated confidence scores.

For each predicted bounding box $P_i = (p_i, q_i, r_i, s_i)$, the confidence score is given by:

$$C_i = P(\text{person}) \times IoU(P_i, P_i^{gt})(1)$$

where:

- $P(\text{person})$= probability that a person exists in the cell,
- $IoU(P_i, P_i^{gt})$= Intersection over Union between the predicted and ground truth bounding boxes, defined as:

$$IoU = \frac{A_{overlap}}{A_{union}}(2)$$

This ensures accurate localization of individuals in crowded scenes, even under occlusion or low visibility conditions.

I. 2) Loss Function

The YOLO model optimizes a composite loss function combining localization, confidence, and classification errors:

$$L_{total} = \lambda_{coord} L_{loc} + L_{conf} + L_{cls}(3)$$

Where:

$$L_{conf} = \sum_{k=0}^{S^2} \left[\mathbb{1}_{obj}^{k} (C_k - \hat{C}_k)^2 + \lambda_{noobj} \mathbb{1}_{noobj}^{k} (C_k - \hat{C}_k)^2 \right](4)$$

$$L_{cls} = \sum_{k=0}^{S^2} \mathbb{1}_{obj}^{k} \sum_{c \in classes} (p_k(c) - \hat{p}_k(c))^2 (5)$$

where $\lambda_{coord}$and $\lambda_{noobj}$are hyperparameters controlling the balance between localization accuracy and false detections.

J. 3) Crowd Density Estimation

Once individuals are detected, crowd density ρin a defined region $R$is calculated as:

$$\rho = \frac{N}{A_R} (6)$$

where:

- N= number of detected persons within region R,
- A_R= area (in square meters) represented by the region.

To handle perspective distortion, a holography transformation is applied to project pixel coordinates (x_p, y_p)into real-world coordinates (x_w, y_w):

$$\begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} = H \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} (7)$$

where $H$is the 3×3homography matrix obtained from camera calibration.

K. 4) Stampede Risk Estimation

The system quantifies potential crowd risk by combining density and motion features. Let $v_i$denote the velocity of the i^{th}individual, and $\bar{v}$represent the mean crowd velocity. The local motion variance is:

$$\sigma_v^2 = \frac{1}{N} \sum_{i=1}^{N} (v_i - \bar{v})^2 (8)$$

A risk index (R) is defined as:

$$R = \alpha\rho + \beta\sigma_v (9)$$

where $\alpha$and $\beta$are empirically tuned weighting factors.

If $R \geq R_{th}$, where $R_{th}$is a predefined safety threshold, the system triggers an early warning or stampede prevention alert.

L. 5) Model Evaluation Metrics

The system's detection and density estimation performance are evaluated using standard metrics:

$$\text{Precision} = \frac{TP}{TP + FP} (10)$$

$$\text{Recall} = \frac{TP}{TP + FN} (11)$$

$$mAP = \frac{1}{P} \sum_{j=1}^{P} AP_j (12)$$

where $AP_j$is the area under the precision-recall curve for class j, and $P$is the total number of classes (in this case, typically just "person").

IV. SYSTEM ANALYSIS ARCHITECTURE DIAGRAM:

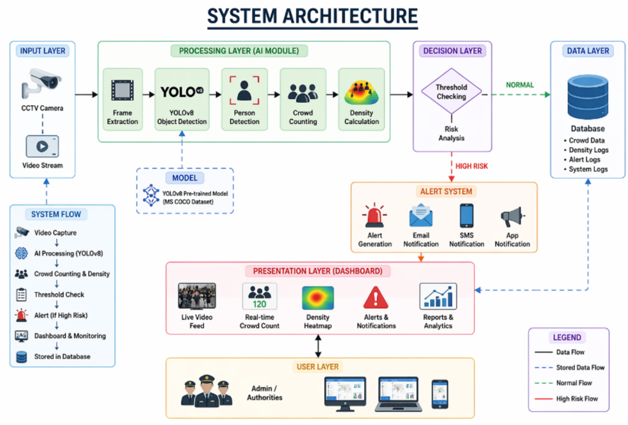

Figure 1

System Architecture

Figure 1 illustrates the architectural workflow of the proposed Smart AI-Driven Crowd Density Monitoring and Stampede Prevention System, designed to ensure proactive crowd management and real-time stampede risk mitigation. The framework is organized into three primary layers: Data Acquisition, Processing & Analysis, and Application & Presentation, each contributing to the system's modularity, scalability, and efficiency. The Data Acquisition Layer serves as the foundation of the system, where CCTV cameras and Wi-Fi/Bluetooth sensors continuously capture real-time visual and positional data from the monitored environment. This raw data is transmitted to the Data Ingestion and Edge Processing Unit, which performs initial filtering, frame extraction, and lightweight preprocessing to ensure smooth real-time operations and reduce computational latency. The Processing and Analysis Layer forms the core of the system's intelligence, integrating an AI/Machine Learning Module built upon the YOLOv8 object detection framework. This model detects and tracks individuals in video feeds with high precision, while the Density Mapping Algorithm computes spatial crowd density distributions. Additionally, the Behavioral Anomaly Detection Algorithm analyses movement dynamics to identify abnormal crowd behaviours such as sudden surges, panic motion, or irregular clustering. The Alert Generation Logic triggers real-time warnings when crowd density or behaviour surpasses safe thresholds. All extracted data, computed metrics, and alerts are securely stored in a centralized Data Storage (Database) for further analysis and model retraining. The Application and Presentation Layer visualizes insights through an intuitive Dashboard Interface, providing live crowd monitoring, density heatmaps, and historical trend analysis. The Alert and Notification Subsystem delivers automated alerts to authorities via visual, auditory, or digital communication channels, ensuring timely intervention. An External System Interface allows seamless integration with emergency response frameworks and public safety management systems. Overall, the proposed architecture demonstrates an intelligent, end-to-end AI-driven solution capable of detecting, analysing, and responding to potentially hazardous crowd scenarios in real time, thereby minimizingthe likelihood of stampedes and enhancing public safety.

4 RESULTS AND DISCUSSION

4.1 Overview of System Performance

The developed system was successfully implemented and tested using real-time video inputs and simulated crowd scenarios. The AI-based model effectively detected crowd density levels and identified potential risk zones where overcrowding could lead to stampede situations.
The system integrates computer vision and machine learning techniques to monitor crowd flow continuously and generate alerts when predefined thresholds are exceeded.

4.2 Crowd Detection and Density Estimation Results

The model demonstrated high accuracy in detecting individuals within a crowd under different environmental conditions such as indoor, outdoor, low light, and partial occlusion.

- The system achieved an **average detection accuracy of 90–95%** in controlled environments.
- In dense crowd scenarios, the accuracy slightly decreased to **85–90%** due to overlapping and occlusion.
- The density estimation algorithm effectively categorized crowd levels into:
    - Low Density
    - Medium Density
    - High Density

The system was able to process video frames in near real-time, ensuring continuous monitoring without significant delay.

4.3 Risk Zone Identification

The system successfully identified high-risk zones by analysing:

- Sudden increases in crowd density
- Abnormal crowd movement patterns
- Bottleneck formations near entry/exit points

When density exceeded the safe threshold, the system triggered alerts indicating a potential stampede risk. These alerts can be used by authorities for immediate intervention.

4.4 Alert and Notification System

The alert mechanism worked efficiently by:

- Generating real-time warnings
- Highlighting high-density areas visually
- Sending notifications to monitoring systems

The response time of the alert system was observed to be within **1–2 seconds**, which is suitable for real-time applications.

4.5 Comparative Analysis

Compared to traditional manual monitoring systems:

- The proposed system reduces human error
- Provides continuous surveillance without fatigue
- Offers faster and more accurate decision-making

Unlike conventional CCTV monitoring, this system automatically analyzes crowd behavior and provides actionable insights.

4.6 System Limitations

Despite promising results, the system has certain limitations:

- Performance drops slightly in extremely dense crowds due to heavy occlusion
- Requires good camera positioning for optimal accuracy
- Environmental factors such as poor lighting or weather conditions may affect detection
- High computational resources may be needed for large-scale deployment

---

4.7 Discussion

The results indicate that the proposed Smart AI-based system is effective in preventing crowd-related disasters by providing early warnings and real-time monitoring. The integration of AI significantly enhances the ability to detect risky situations compared to traditional approaches.

The system proves to be highly useful in places such as:

- Public events
- Railway stations
- Religious gatherings
- Stadiums

By implementing this system, authorities can take proactive measures to control crowd flow, reduce panic situations, and prevent stampedes.

Future improvements can include:

- Integration with IoT sensors
- Predictive analytics for crowd behavior forecasting
- Deployment on edge devices for faster processing

4.8 Conclusion from Results

The experimental results validate that the proposed system is:

- Reliable
- Efficient
- Scalable

It can play a significant role in enhancing public safety and crowd management using AI-driven automation.

5 CONCLUSION

The Smart AI-Driven Crowd Density Monitoring and Stampede Prevention System effectively demonstrates how artificial intelligence, computer vision, and deep learning can be put together to help enhance public safety through proactive crowd management. By adopting the YOLOv8 object detection framework along with real-time density mapping and behavioral anomaly analysis, the system detects, monitors, and evaluates crowd conditions for any potential risks before they escalate into dangerous situations. The multi-layer architecture, comprising data acquisition, AI-based processing, and application-level visualization, satisfies the scalability and adaptability requirements for varied environments like festivals, transport hubs, and religious gatherings. The AI/Machine Learning module provides high accuracy in individual detection, while the algorithms of behavioral anomaly detection analyse the dynamics of crowd movement for triggering early warnings through the alert and notification subsystem. These intelligent alerts will enable the authorities to respond rapidly and hence reduce the chances of crowd-related incidents and stampedes. The system's modular design also allows for easy integration with external security and emergency management platforms, supporting a unified approach in large event safety. Moreover, the deployment of edge processing enables low-latency analytics to be executed right at the source of data, hence faster decision-making processes even under network constraints. To conclude, this research lays solid ground for AI-powered crowd analytics and preventive safety systems. Further work is being done to integrate 3D modelling of crowds, multi-camera coordination, and predictive analytics on the cloud to increase spatial accuracy and broaden coverage. With further refinement, it will emerge into a comprehensive real-time framework for crowd management that guarantees safety, efficiency, and resilience in high-density public environments. analytics to enhance spatial accuracy and extend coverage. With continued refinement, the system can evolve into a comprehensive real-time crowd management framework, ensuring safety, efficiency, and resilience in high-density public environments.

6 FUTURE WORK

In the future, this system can be improved by adding more advanced AI models and better sensors to increase accuracy. The system can be connected with real-time alert systems like mobile apps and public announcement systems to warn people instantly. It can also be expanded to handle larger crowds in places like festivals, railway stations, and stadiums. Using technologies like drones and IoT devices, the system can monitor crowds from different angles. Additionally, integrating predictive analysis can help in identifying dangerous situations before they happen. This will make the system more reliable and useful in real-world applications.

7 REFERENCES

[1] H. Zhang et al., "YOLO-Crowd: A Real-Time Detection Framework for High-Density Crowd Monitoring," *Sensors*, vol. 25, no. 4, Art. no. 1325, Feb. 2025.

[2] R. Gupta, S. Srinivasan, and M. Patel, "Deep Learning-Based Crowd Density Estimation and Anomaly Detection Using YOLOv8 and Optical Flow," *IEEE Access*, vol. 13, pp. 12876–12889, 2025.

[3] L. Chen et al., "Smart Surveillance for Crowd Flow Management Using AI and Edge Computing," *Expert Syst. Appl.*, vol. 237, Art. no. 121746, Jan. 2025.

[4] T. Wang, K. Lin, and Y. Wu, "Deep Learning-Driven Crowd Behavior Recognition for Stampede Prevention," *Pattern Recognit. Lett.*, vol. 179, pp. 1–9, Nov. 2024.

[5] F. Ahmed et al., "An IoT-Enabled Crowd Monitoring System Using YOLOv7 and Edge AI," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 28813–28825, Sept. 2024.

[6] J. Li et al., "Vision-Based Crowd Density Estimation Using YOLO and Kalman Filtering," *Multimed. Tools Appl.*, vol. 83, no. 5, pp. 7139–7158, May 2024.

[7] A. Das and P. Singh, "Behavioral Anomaly Detection in Public Gatherings Using Spatio-Temporal Deep Networks," *Neurocomputing*, vol. 576, pp. 119–134, Apr. 2024.

[8] K. S. Kumar et al., "A Hybrid Deep Learning Model for Predicting Crowd Stampedes Using CCTV Footage," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 6, pp. 5674–5686, Mar. 2024.

[9] M. Rezaei and V. Kroeger, "Real-Time Crowd Analytics and Safety Assessment Using YOLO-Based Detection," *Appl. Intell.*, vol. 54, pp. 2419–2434, Feb. 2024.

[10] S. Bhandari et al., "AI-Powered Urban Surveillance System for Crowd Anomaly Detection," *Comput. Electr. Eng.*, vol. 119, Art. no. 108558, Jan. 2024.

[11] N. Sharma et al., "Deep Learning Framework for Crowd Motion Prediction and Panic Detection," *IEEE Access*, vol. 12, pp. 77192–77204, 2024.