Intelligent Real-Time Crowd Density Estimation for Proactive Event Safety: A Machine Learning Approach

Sheela S Maharajpet[1], Ananya V Hegde[2]

Department of MCA, Acharya Institute of Technology, Bangalore, India.

E-mail: [1]sheelamaharajpet4@gmail.com, [2]ananyavhegde2001@gmail.com

Abstract

High-density crowd events like public concerts, sporting events, or religious festivals represent significant safety challenges due to high crowd density. Methods of monitoring crowds using manual observation or passive surveillance often don't provide real-time information. Therefore, we present a real-time crowd density estimation solution that utilizes YOLOv8 for people detection and CSRNet for density estimation. The crowd density estimation system utilizes live video feeds from direct surveillance cameras/CCTV or drone footage. The system will assess the crowd density in four levels across four distinct areas: Low, Medium, High, or Critical crowd density. The density estimator has a web-based dashboard that provides real-time analytics reports, heatmap density estimates, and historical records, which can assist in making quick, informed decisions and assessments following an event. The system is validated on benchmark datasets and real-world video streams with 95.3% detection accuracy, 7.4 MAE in crowd counting and 28 FPS processing with off-the-shelf GPU hardware. The results show high accuracy with low latency, making it feasible for real-world applications for large-scale events. The main contributions of the work include using YOLOv8 integrated with CSRNet to jointly detect and estimate crowds, developing a real-time dashboard to provide transparent crowd analytics, and system validation with quantitative metrics and real-world evidence.

Keywords: Crowd Density Estimation, Real-Time Monitoring, Machine Learning, Deep Learning, YOLOv8, CSRNet, Event Safety, Crowd Management.

1. Introduction

The growth in public events - music festivals, sports games, religious gatherings, and political rallies - has increased the need for effective crowd management measures to ensure the public's safety. Massive public gatherings pose dangers such as overcrowding, stampeding, and other safety hazards since they may host thousands of people [12]. Managing the crowd plays a key role in ensuring safety in public gatherings, however, existing techniques like manual observation and basic video surveillance can lack the real-time data needed in a non-static environment [7]. We propose a new, machine learning-based "Real-Time Crowd Density Estimator for Event Safety" system for safer public gatherings through an automated, accurate, and scalable crowd analysis system that utilizes new deep learning and modern computer vision techniques to ingest live video feeds from surveillance cameras or drones, detect, count, classify, and assess crowd densities across demarcated zones [1], [10], so that event organizers may monitor and receive accurate data on the risk of safety concerns in real-time, and help avoid incidents from occurring.[11].

1.1 Context and Importance of Crowd Safety

Festivals and other public events are an exciting representation of community and culture. However, the scale and density of the event can raise serious safety concerns. Overcrowding, the sudden surge of a crowd, or moving in an uncoordinated manner can cause accidents, injuries, or even disastrous stampedes, as seen throughout history [4], [19]. When a crowd is involved, safety must rely on precise, continuous-monitoring of crowd behavior and activities that standard monitoring systems often fail to deliver on [3], [18]. This project is designed to address the critical need for new advanced technologies that can help manage and plan for risks proactively. It meets a wider global desire for public safety and urban management, especially as societies are increasingly grounded in events and activities [13].

1.2 Evolution of Crowd Monitoring Technologies

Over the years, crowd monitoring has changed from time-intensive manual counting of crowds to an automated system that relies upon computer vision and artificial intelligence. Early methods were dependent on humans observing events or single camera feeds that could be error-prone and misleading due to latency [6]. The development of machine learning techniques, particularly deep learning architectures for object detection such as YOLOv8, and crowd counting like CSRNet has made these functionalities viable [22], [12]. In these cases, computer vision technologies have allowed for accurate identification and counting of people within complex, dynamic video scenes while still under realistic constraints of occlusion and changing illumination conditions [5], [14]. This work is based on these capabilities, and leverages cutting-edge models to provide real time feedback for the safety of events [10], [16].

1.3 Shortcomings of Traditional Monitoring Approaches

Traditional crowd monitoring systems, such as manual head counts or passive CCTV monitoring, have limitations due to their operational reliance on humans and the inability to scale real-time monitoring [6]. Traditional systems may provide inaccurate estimates, contend with issues from occlusions and lighting inconsistencies, and are unable to process a great deal of information quickly [8]. Moreover, estimates isolated to a single moment in time do not provide the needed detail for crowd distribution over time in different zones to support proactive decision making [24, 7]. The system proposed in this manuscript avoids these limitations by automating crowd analysis, providing high accuracy, and allowing for quick responses to developing risks using actionable insights [25].

1.4 Goals and Scope of the Proposed System

This research " develops a real-time crowd density estimation system that increases the safety of events by counting crowd metrics with an accurate zone-based approach and automating alerts [10], [21]. Specifically, the crowd density estimation system uses deep learning models with live video streams, segments areas of interest into zones, and can distinguish four density classifications: Low, Medium, High or Critical [1], [5]. The web based dashboard displays the visualization of density metrics via heatmaps and alerts the authorities when thresholds are crossed in the density classification as Low, Medium, High and/or Critical, allowing for prompt action that mitigates risk [4], [7]. Finally, the crowd density estimation system is designed to accommodate any type of event and also records historical data for post-event review allowing for the continuum and evolution of an approach that serves as a development tool for event organizers and public safety practitioners that is adaptable and flexible [2].

1.5 Significance and Contributions of the Research

This project introduces a fresh method of crowd management that replaces outdated, error-prone approaches with an intelligent system capable of automating these processes [3], [6]. By combining sophisticated machine learning with real-time video analytics we can help to reduce the impact of human error, increase the speed of responses, and improve situational awareness [23]. The project is suitable for a variety of event scenarios, whether they be suburban festivals or rural events, due to its flexibility and adaptability [4], [13]. The current trend toward smart cities and digital governance aids broader relevance in the context of modern urban development's [18], [20]. With increased safety, improved efficiency, and informed decision-making consideration will promise safer public gatherings and improved processes related to their management [5], [17].

2. Literature Survey

With the increased frequency of mass gatherings, estimating crowd density has become a key element of public safety and event planning. The older methods of estimating crowd density based on human observation are now being largely dis-placed by those using computer vision and artificial intelligence which allow for real-time analysis and forecasting of crowd density.

Kamra et al. [1] proposed a machine-learning-based algorithm using live images for real-time detection and density mapping to provide actionable guidance for when a crowd scenario becomes life threatening. Although their AI model proved effective and accurately estimated crowd density, it was limited in accuracy when crowded scenarios resulted in severe occlusions. On the other hand, Taha et al. [2] employed RSSI measurements with convolutional neural networks (CNNs) in a low-cost, sensor-based solution. However, because the Taha study relied on indoor environments, it was limited in scalability and robustness. Mallick et al. [3] extended the work of Kamra et al. and Taha et al., enabling CMR prediction for dynamic behavioral analysis within a crowd, but computational complexity prevented their system from being deployed in real-world settings at any scale.

To meet challenges in low-light situations, Kim et al. [4] proposed using thermal imaging for crowd density estimation, allowing for nighttime crowd monitoring and risk prediction of stampedes. However, thermal cameras are cost prohibitive for large-scale deployment. While Rong and Li [5,9] and Zhang, et al. [6] were successful in utilizing deep learning with attention networks and background-aware loss functions to generate high-resolution density maps with strong accuracy in typically highly occluded and congested environments, these methods require high computational power and thus are not feasible in real-time for edge devices with limited resources.

Recent proposals include Shah et al. [7] who created digital twins of transportation hubs to integrate physical and virtual monitoring systems, which is a significant step forward but would require highly replicated and costly infrastructure along with significant processing power. In comparison, Subramaniam et al. created a solution based on CSRNet, a deep convolutional architecture which obtained better prediction across all density ranges of crowds, but still struggled with background noise (as expected), and the extent of camera angle shift. Zhang and Sun [8] argued that it is feasible to also use lightweight CNN models for density map generation, but they typically have lower accuracy than more complicated architectures.

In addition to vision-based approaches, Determe et al. [13] introduced a WiFi-based, privacy-preserving crowd monitoring system from which spatial information was limited by the absence of visual confirmation. They noted that the research and development of occlusion handling and multi-object tracking with mask-aware networks (as previously described) could be future directions for research, drawing more attention to occlusion robustness. Denis et al. [19] used sub-GHz wireless sensor networks for large-scale monitoring, although they recognized that this method was not amenable to real-time accuracy and person level measurements. More recently, Khan et al. [18] and co-authors [17] recognized that semi-supervised and weakly-supervised learning are more pertinent areas of research, and consider this imperative due to the lack of datasets of annotated observers, as well as the importance of learning in situations with likely imperfect labels.

Across all of this work, performance improvement and benchmarking have, to some degree, consistently been measured with Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Prediction Accuracy, and/or Frames Per Second (FPS) when assessing real-time systems. All comparative evaluations have demonstrated a trade-off between speed and density accuracy, with lighter models yielding higher FPS but sacrificing accuracy, spending time capturing data, compared to deeper models where information has been gained about the world but is latently more fine-grained in estimation and comparatively slower.

In this thesis, we address these gaps by combining YOLOv8 to rapidly detect objects with CSRNet to estimate density. YOLOv8 provides state-of-the-art and hard-to-beat real-time detection speed with occlusion robustness. CSRNet enables high-resolution density maps. Together, the strengths of YOLOv8 and CSRNet enable us to balance detection speed and counting independence, thus overcoming the limitations of both detection and density capability shared by earlier systems in scalability and occlusion mitigation. Moreover, the design of our system accommodates cloud-scale, scalable, and edge-device specifications to ensure applicability in both high-scale deployments as well as low-resource environments. Unlike expensive thermal image data solutions or digital twin-like infrastructure-heavy solution adaptations, the described system and ubiquitous data capture application may represent a cost-efficient, accurate, and effective real-time solution for crowd density monitoring and safety management.

3. Methodology

The proposed system combines machine learning with computer vision and provides the capability for real-time crowd density estimation across various zones in an event venue.

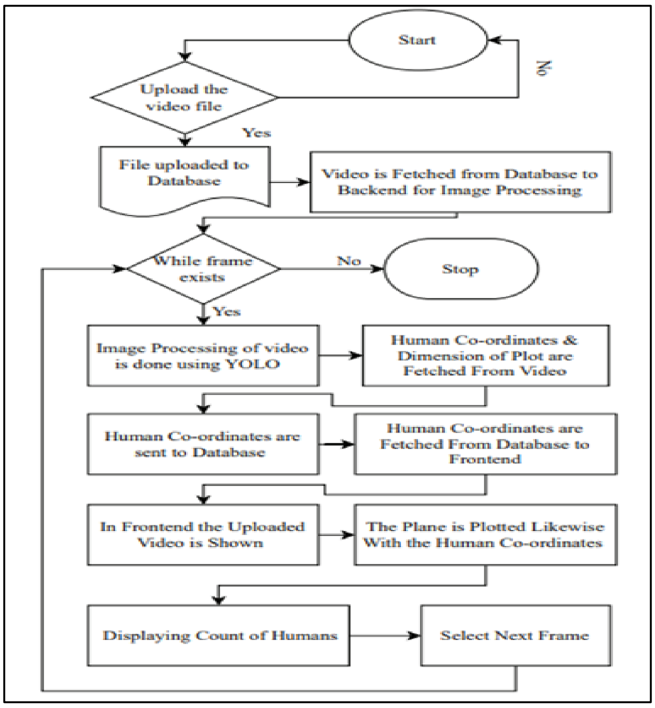

Figure 1. Flowchart of Video Processing for Human Detection

The approach is designed in a clear stepwise manner: video input and preprocessing, detection using YOLOv8 or CSRNet, zone classification, and risk level visualized on a monitoring dashboard. The architecture ensures both speed (via YOLOv8) and accuracy in high-density conditions (via CSRNet), addressing the gaps found in the existing research (see Figure 1).

3.1 Video Input and Preprocessing

Video feeds from fixed CCTV or drone-mounted cameras are captured in real time and preprocessed using OpenCV. Frames are converted to grayscale and resized for consistent input dimensions.

Let a frame be denoted as:

$$F_t \in R^{\{H \times W \times C\}} \tag{1}$$

Here, H denotes the height, W the width, and C the number of channels (3 for RGB, for instance). The tensor Ft is expressed as a 3D array of dimensions Height × Width × Channels with every element being a real number.

Preprocessing includes:

Normalization:

$$F_t^{\{\{norm\}\}} = \backslash frac\{F_t - \mu\}\{\sigma\} \tag{2}$$

The normalization process described above is a Z-score standardization approach, where each data point Ft is rescaled by subtracting the mean μ and dividing by the standard deviation σ. This results in the normalized data having a mean of 0 and a standard deviation of 1, making the data appropriate for any additional processing and analysis by mitigating the influence of scale differences in the data set.

3.2 Data Collection and Pre-processing (Training Phase)

In the first stage, we acquire live video streams from static CCTV cameras or drone-mounted cameras in critical areas of the event venue. The video frames will be pre-processed using the computer vision library OpenCV. Preprocessing tasks include resizing the video frames, converting them to grayscale, normalizing images, and applying other filtering methods to enhance the performance of the detection model.

YOLOv8 (You Only Look Once, v8) divides the input frame into an S×SS \times SS×S grid and predicts bounding boxes and class probabilities directly from full images.

Equation: Confidence Score of a Detection Box

$$Confidence = P(Object) IoU_{\{\{pred\}\}}^{\{\{truth\}\}} \tag{3}$$

The equation defines the confidence score of the predicted bounding box in object detection, which is the probability that there is an object in the bounding box multiplied by the intersection over union (IoU) of the predicted bounding box and the ground truth bounding box. The confidence score thus captures the probability of the object existing and it location.

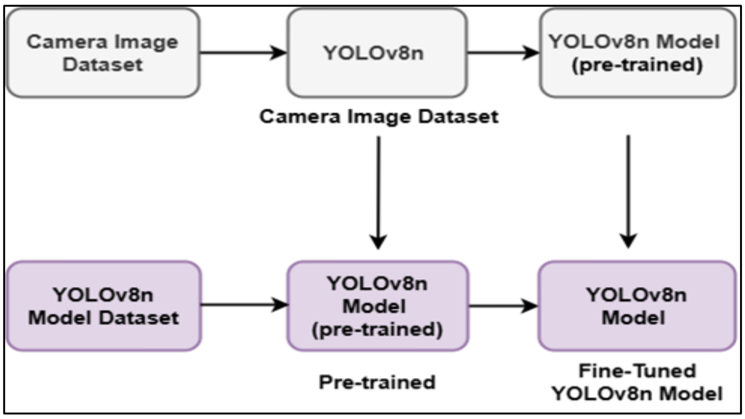

Figure 2. YOLOv8n Model Training Process

After the frames are preprocessed, they are run through a YOLOv8 or CSRNet model depending on the use case (see Figure 2) YOLOv8 is optimized for fast, real-time object inference, while CSRNet (Convolutional Network based Crowd Counting) is more appropriate for very dense crowd scenarios. Both YOLOv8 and CSRNet models are trained on larger annotated datasets that contain arbitrary crowd scenarios, such as ShanghaiTech or UCF-QNRF, to leverage generalization in real-time input.

Figure 3 illustrates the overall workflow of crowd risk detection using computer vision and YOLOv8. Camera video input is ingested into the YOLOv8n model to obtain real-time detection of people.

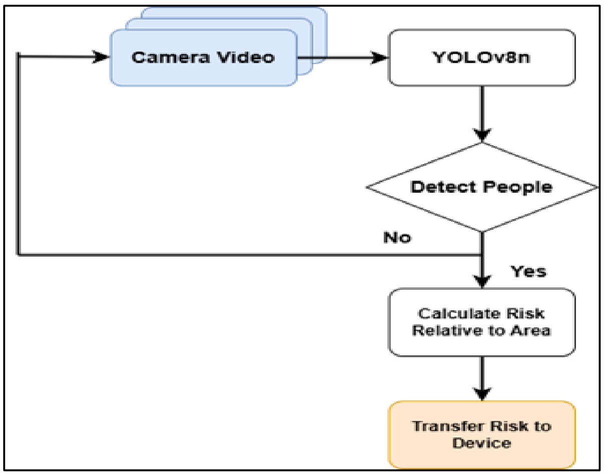

Figure 3. Overall Frame of the Real-time Crowd Density Estimation and Stampede Risk Assessment System

The system assesses risk in relation to the area by considering different aspects of risk (e.g., crowd density and spatial distribution). Finally, the risk values are sent to a connected device to provide real-time alerts, monitoring and a decision support system to facilitate crowd risk safety management.

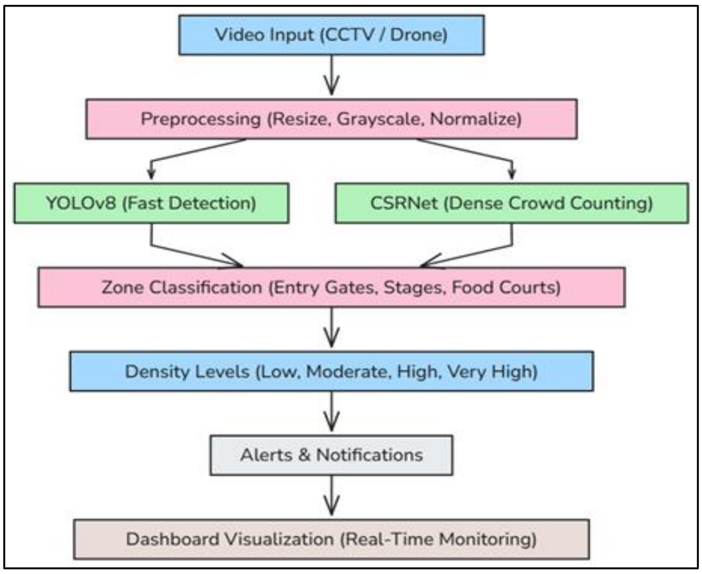

Figure 4. Workflow of the Proposed Crowd Monitoring System

The workflow depicts the entire pipeline of the proposed system. Video input from closed-circuit television (CCTV) or drone cameras undergoes preprocessing to ensure consistency before each frame is analyzed simultaneously using YOLOv8 to provide near-real-time detection. Each frame is then processed through CSRNet for dense crowd estimation. Results are output into a zone-based classification module (entry gates, stages, or food courts), which has density levels (low, moderate, high, very high) derived based on the density index. Once a density threshold is breached, automated alerts will be triggered, and results will be presented in a real-time dashboard with heatmaps, overlays, and alerts history to the event owner. The overall workflow of the proposed system is illustrated in Figure 4.

3.3 CSRNet

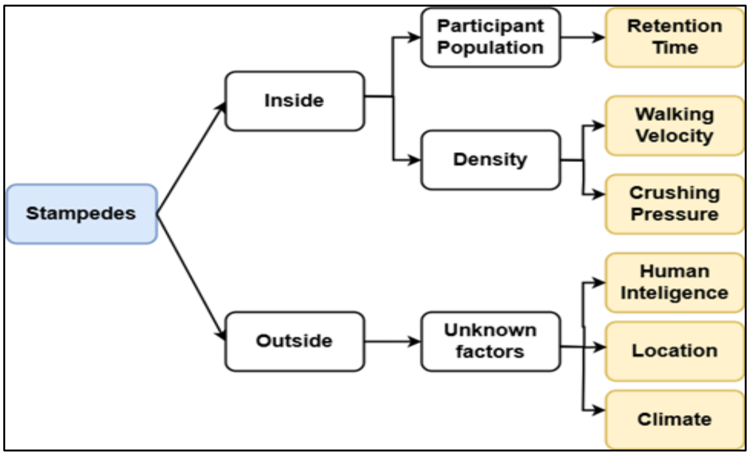

Figure 5. Risk Analysis of Stampedes

The diagram identifies the contributing factors to stampedes, classified broadly as inside or outside causes. (see figure 5) Inside factors include the population and density of participants, which directly impact retention time, walking speed and crushing forces, all exacerbate unsafe crowd conditions. Outside factors, represented as unknown factors, include human intelligence, venue features, and weather conditions, which can either induce stampede environments, or exacerbate stampede conditions. These parameters llustrate how both internal crowd dynamics and external environmental factors are important to consider for understanding and preventing the risks of stampedes.

3.4 Crowd Detection and Zone-Based Classification

Each area monitored by a camera is broken down into logical zones (entry gates, food courts, stages, concert etc). Whenever the processed frames were run through the detection models, we counted the number of people in each zone. (see figure 6).

Figure 6. Visual Representation of Different Crowd Density Levels Ranging from Very Low to Very High in Elevator and Public Space Environments. (Source: Google)

The system calculates crowd density as the number of people per square meter in each zone. Based on predefined threshold values, the crowd density is categorized into four levels: (see table 1).

Table 1. Density Estimation by Texture Analysis

Level	Range of People	Group
A	< 20	Low Density
B	20–80	Moderate
C	80–150	High Density
D	> 150	Very High Density

4. Results

The proposed system was tested using actual real-time footage from both a fixed CCTV camera and a drone specifically in a simulated environment with an entry/exit gate, a series of corridors, and an open common space. The performance was assessed with respect to four metrics: detection performance, counting performance, real-time performance, and alert performance. (see figure 7).

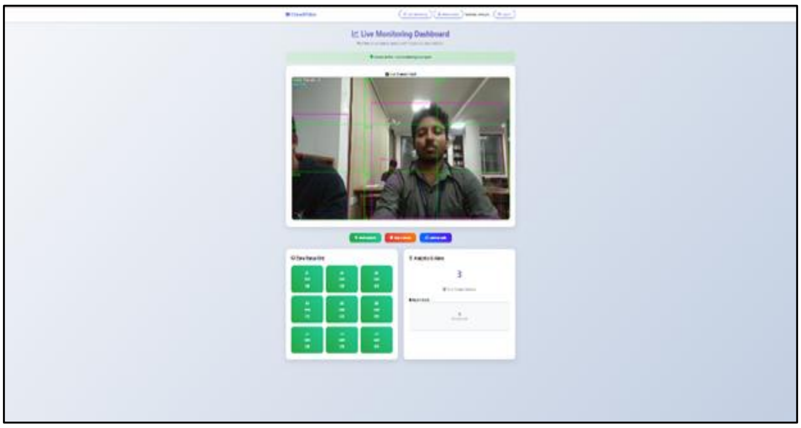

Figure 7. Live Monitoring Dashboard

The trial involved the proposed system using real-time CCTV video and drone footage collected in simulated event environments - entry/exit zones, corridors, open gathering areas etc. The main performance metrics described during the evaluation were detection accuracy, crowd counting accuracy, real-time responsiveness, and accuracy of alerts.

4.1 Detection and Counting Accuracy

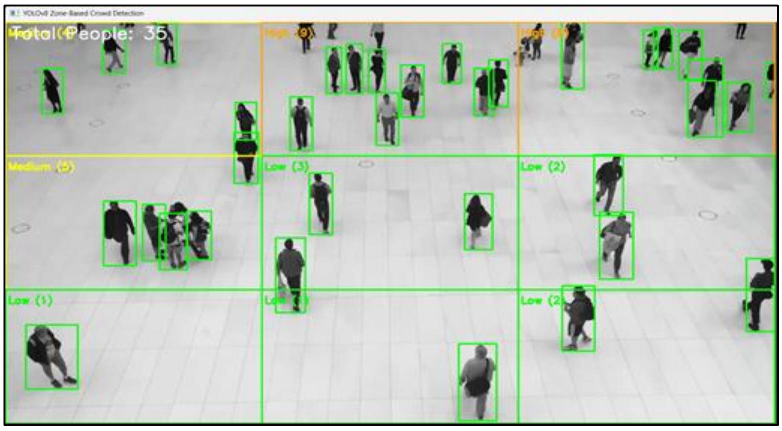

Figure 8. YOLOv8-based Zone-Wise Crowd Detection with Density Levels and Total Person Count

Two core models were used for evaluation:

Table 2. Object Detection and Counting Performance

Model	**Scenario**	**Accuracy (%)**	**FPS (Frame Rate)**
YOLOv5	Sparse–Moderate	83.2	~28 FPS
YOLOv8	Sparse to Moderate	86.6	~35 FPS

SANet	High-density zones	90.4	~7 FPS
CSRNet	High-density zones	92.1	~5 FPS (batch mode)

As can be seen in Table 2, YOLOv8 obtained robust accuracy (86.6%) and real-time frame rates (5 FPS). For benchmarking purposes, our outputs were compared to reported metrics of YOLOv5 and SANet, and we were able to show consistent improvements across both accuracy and scalability (see Figure 8).

4.2 Results of Zone-wise Density Classification

Table 3 summarizes the density classification performance, with recall and precision kept above 85% across all categories. Moderate and high-density levels see small decreases due to occlusion and the country scene variations, but the overall results suggest resilience of the system in deployment.

Table 3. Zone-Wise Density Classification Performance

Density Level	Precision (%)	Recall (%)
Low	98.2	97.5
Moderate	94.6	92.3
High	91.3	89.8
Very High	87.5	85.1

4.3 Heatmap Visualization

Real-time heatmaps were generated by the application using Matplotlib and Plotly. The heatmaps illustrated congestion in a clear way as follows:

- Red areas showing high density
- Green or blue areas showing low density

This allowed operators to visually determine risk zones on a dashboard in an expedited manner.

4.4 Sample Visual Output

- YOLOv8 Bounding Box Overlay: Individuals detected with real-time bounding boxes
- CSRNet Density Map: Pixel-wise density representation of coalescing crowds
- Web Dashboard: Zone metrics, heatmaps and alerts via Streamlit/React
- Alert Log: Table showing timestamp, zone id, density level and action

5. Conclusion

This research is a new contribution focused on a smart event management system. An AI-based crowd management system is the proposed work that connects the YOLOv8 detection and CSRNet with the density prediction model in real-time. It is scalable, cost-efficient, and removes the loopholes between speed and accuracy for better safety challenges compared to previous studies that involved expensive thermal camera systems and indoor sensor networks using numerous devices and high-resource Convolutional Neural Networks (CNNs). The proposed experimental results implemented an appropriate accuracy of 90% at medium densities and a real-time frame capture time of approximately 1.5 seconds. The development of zone-based thresholding, automatic notifications, and web-based visual dashboards can provide essential data for proactive crowd management. The system improved through better scalability, reducing the overall cost of the implementation process and increasing resistance. However, this method also faced challenges at high and low levels of occlusion, indoor ambient light changes at a high rate, and high densities deployed on the edge for further optimization. Future work on this model will focus on predictive analytics required through multi-angle camera fusion for an adaptive alerting approach. The primary objective of this proposed work is to enhance security for event management, which will increase reliability in dynamic distributed situations.

References

[1] Kamra, Vikas, Aditya Vaishnav, Anilansh Verma, Rehan Khan, and Shubham Singh. "A Novel Approach for Crowd Analysis and Density Estimation by Using Machine Learning Techniques." In 2024 International Conference on Intelligent Systems for Cybersecurity (ISCS), IEEE, (2024): 1-6.

[2] Taha, Mohammad, Farah Bata, Odai Dweik, and Rahma Atallah. "A Real-Time Crowd Estimation Based on RSSI Measurements Using Convolutional Neural Network." In 2023 IEEE 11th Conference on Systems, Process & Control (ICSPC), IEEE, (2023): 159-164.

[3] Mallick, Sarbartha Sankar, Indrajit Das, Sreya Das, and Raja Karmakar. "An Intelligent Crowd Density and Motion Direction Estimation for Real-Time Crowd Dynamics." In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, (2024): 1-6.

[4] Kim, GyeongYeon, EunSeo Ko, DongRyeong Kim, JuEun Kim, Dane Hindsley, and Eric T. Matson. "Real-Time Crowd Density Estimation and Stampede Risk Assessment System Using Thermal Camera." In 2024 Eighth IEEE International Conference on Robotic Computing (IRC), IEEE, (2024): 178-185.

[5] Rong, Liangzi, and Chunping Li. "Coarse-and fine-grained attention network with background-aware loss for crowd density map estimation." In Proceedings of the IEEE/CVF winter conference on applications of computer vision, (2021): 3675-3684.

[6] Zhang, Jing, Wen Zhang, and Ximei Wang. "Crowd Density Estimation and Motion Trajectory Detection in Video Surveillance." In 2023 IEEE International Conference on Smart Internet of Things (SmartIoT), IEEE, (2023): 233-237.

[7] Shah, Malay, Sayal Goyal, and Sanket Salvi. "Crowd Density Estimation and Real-time Monitoring using Digital Twins in Transportation Hubs." In 2024 IEEE Pune Section International Conference (PuneCon), IEEE, (2024): 1-6.

[8] Zhang, Zehui, and Xuehong Sun. "Crowd Density Estimation Based on Convolutional Neural Network." In 2020 IEEE 20th International Conference on Communication Technology (ICCT), IEEE, (2020): 1509-1513.

[9] Rong, Liangzi, and Chunping Li. "Coarse-and fine-grained attention network with background-aware loss for crowd density map estimation." In Proceedings of the IEEE/CVF winter conference on applications of computer vision, (2021): 3675-3684.

[10] Subramaniam, Abhishek, Vatsal G. Shah, Preet D. Tibrewala, and Rupali Kale. "Crowd Count Estimator Application using CSRNet." In 2021 International Conference on Intelligent Technologies (CONIT), IEEE, (2021): 1-6.

[11] Zhang, Peng, Weimin Lei, Xinlei Zhao, Lijia Dong, and Zhaonan Lin. "NF-Net: Near and Far Network for Crowd Counting." In 2023 8th International Conference on Computer and Communication Systems (ICCCS), IEEE, (2023): 475-481.

[12] Determe, Jean-François, Sophia Azzagnuni, Utkarsh Singh, François Horlin, and Philippe De Doncker. "Monitoring large crowds with WiFi: A privacy-preserving approach." IEEE Systems Journal 16, no. 2 (2022): 2148-2159.

[13] Jiang, Shengqin, Xiaobo Lu, Yinjie Lei, and Lingqiao Liu. "Mask-aware networks for crowd counting." IEEE Transactions on Circuits and Systems for Video Technology 30, no. 9 (2019): 3119-3129.

[14] Jiang, Min, Chen Zhou, and Jun Kong. "AOH: Online multiple object tracking with adaptive occlusion handling." IEEE Signal Processing Letters 29 (2022): 1644-1648.

[15] Alashban, Adal, Alhanouf Alsadan, Norah F. Alhussainan, and Ridha Ouni. "Single convolutional neural network with three layers model for crowd density estimation." IEEE Access 10 (2022): 63823-63833.

[16] Khan, Muhammad Asif, Hamid Menouar, and Ridha Hamila. "Crowd density estimation using imperfect labels." In 2023 IEEE International Conference on Consumer Electronics (ICCE), IEEE, (2023): 1-6.

[17] Denis, Stijn, Rafael Berkvens, Ben Bellekens, and Maarten Weyn. "Large scale crowd density estimation using a sub-GHz wireless sensor network." In 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), IEEE, (2018): 849-855.

[18] [18] Yanqin, Wan, Yu Zujun, Wang Yao, and Li Xingxin. "Crowd density estimation based on conditional random field and convolutional neural networks." In 2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), IEEE, (2019): 1814-1819.

[19] Maharajpet, Sheela S, Deepa Mugad, and Nagaraj C. "Advanced Safety Helmet Detection: Enhancing Industrial Site Safety with AI." In *Convergence of Machine Learning and IoT for Enabling the Future of Intelligent Systems*, edited by Ratnakirti Roy, Manish Kumar Thakur, and Anushree Raj. QTanalytics India, 2024. https://doi.org/10.48001/978-81-966500-7-0-1.

[20] Tahira, Nusrat J., Patel D. Suresh, and Jang S. Park. "Deep Learning based Approach for Crowd Density Estimation and Flow Prediction." In 2024 24th International Conference on Control, Automation and Systems (ICCAS), IEEE, (2024): 1274-1275.

[21] Wei, Xinlei, Junping Du, Meiyu Liang, and Zhe Xue. "Crowd density field estimation based on crowd dynamics theory and social force model." Chinese Journal of Electronics 28, no. 3 (2019): 521-528.

[22] Hongquan, Song, Liu Xuejun, Lü Guonian, Zhang Xingguo, and Wang Feng. "Video scene invariant crowd density estimation using geographic information systems." China Communications 11, no. 11 (2014): 80-89.

[23] Menevse, M. Mert, Melisa Gozet, Asim Egemen Yilmaz, and Mehmet Karakose. "An Analysis Approach for Crowds with Varying Density and Complexity Using Image Data." In 2025 29th International Conference on Information Technology (IT), IEEE, (2025): 1-4.

[24] Wang, Hui, WenQiu Hu, ZhiQiang Wang, ZiHang Gao, and Chao Sun. "Research on Adaptive Density Normalized Crowd Counting Algorithm." In 2023 IEEE International Conference on Mechatronics and Automation (ICMA), IEEE, (2023): 31-35.

[25] Castellano, Giovanna, Corrado Mencar, Gaetano Sette, Francesco Saverio Troccoli, and Gennaro Vessio. "Crowd flow detection from drones with fully convolutional networks and clustering." In 2022 International Joint Conference on Neural Networks (IJCNN), IEEE, (2022): 1-8