WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Real-Time Public Safety and Crowd Management

Sirikonda Akhila[1], Srinivas Rao Kunchala [2]

[1]*Student, School of Informatics, Department of MCA, Aurora Deemed University, Hyderabad*
[2]*Assosciate Professor, School of Informatics, Department of MCA, Aurora Deemed University, Hyderabad*
akhilachowdarysiirkonda21@gmail.com, 2Srinivasrao@aurora.edu.in

ABSTRACT –

Urban public safety has emerged as a priority with the hasty growth in population, traffic jams, and unexpected emergencies. Conventional surveillance techniques like manual observation using CCTV cameras do not work, are susceptible to human mistake, and cannot cope with high-volume real-time scenarios. To counter these issues, Artificial Intelligence (AI) provides state-of-the-art solutions integrating computer vision, deep learning, and real-time data processing to effectively handle crowds and maintain public safety.

This research is cantered on the use of AI-driven crowd analysis and detection software for street, public event, and city real-time monitoring. By employing Convolutional Neural Networks (CNNs) and object detection models like YOLO (You Only Look Once), the system is able to identify people in real-time video feeds automatically, estimate crowd density levels, and mark out-of-the-norm crowd activity like instantaneous congestion, violent outbursts, or panic unrest. In addition, utilization of global density estimation methods enhances accuracy for densely congested situations where standard detection models do not work.

The suggested method not only displays visual information as bounding boxes and heatmaps but also initiates real-time alarms for police officers and urban planners to prevent possible hazards. Training and testing are conducted on real datasets of road surveillance and crowd formations to attain the system's reliability in changing environments.

Its potential use is in traffic management, disaster evacuation planning, crowd monitoring during festivals or concerts, and crime management. From the optimized detection and management process, this study tells us how AI can expand situation awareness, decision-making, and rapid response in some substantial ways to render cities safe and resilient.

Keywords: Crowd Management, Artificial Intelligence, YOLO, Computer Vision, Public Safety, Real-Time Monitoring, Edge Computing

I. Introduction

High population growth and urbanization have made city roads very crowded and lively. With the millions passing through them daily, visiting public meetings, or visiting public forums, public safety and effective crowd management have emerged as the government and the law enforcement authorities' prime concerns. Surveillance systems that have been based on classical human observation of CCTV feeds are not performing due to their limited scalability, responsiveness, and susceptibility to human-made errors. All these weaknesses foster the need for intelligent, autonomous, and real-time solutions that have the capability to augment situational awareness and decision-making within cities.

Over the past decade, Artificial Intelligence (AI) has been a valuable tool to address these challenges. Leveraging computer vision, deep learning, and edge computing technologies, AI-powered systems have been able to detect, analyse, and predict human behaviour trends in real time. Particularly, object detection methods like YOLO (You Only Look Once) and density-based methods have been very effective in the estimation of crowd density, movement analysis, and threat detection like congestion, stampede, or suspicious behaviour. These systems can process raw video streams and transform them into action and actionable intelligence, which can have the potential to prompt faster response and preventive actions.

Utilization of AI to manage crowds and observation systems for street watch enhances effectiveness and accuracy and offers dynamic and adaptive solutions for differentiated cities. For instance, AI can help road traffic congestion to be managed by the police, support event organizers to monitor crowd movement in public festivities, and allow security guards to detect and react to suspicious actions in real time. In addition, with the inclusion of visualisation tools like heatmaps, density maps, and anomaly detection, decision makers have a greater understanding of the situation than they would otherwise have had.

AI for the Streets, then, is a shift from reactive monitoring to smart, real-time, and proactive public safety systems. Capable of learning like machines, real-time analysis, and visualization, the approach provides the devices for safer, wiser, and more resilient cities to efficiently address the needs of public safety and crowd management today.

II. Literature Review

1) Pre-Deep Learning Computer-Vision Methods

Background subtraction and optical flow were the first methods employed in street CCTV to humans and motion inference. Background models (i.e., Gaussian Mixture Models) remove moving blobs from stationary environments; optical flow fields are composed of motions at the pixel level. The methods are easy and can be executed in real-time but fare badly with illumination changes, camera motion, shadows, and dense occlusion of crowds.

Hand-tuned methods like HOG (Histograms of Oriented Gradients) + SVM are the state of the art for pedestrian detection nowadays. HOG is an edge direction representation in fixed-size windows (e.g., 64×128), and a linear SVM classifies it as "person vs non-person." It performs fine on mid-height standing people in good light and remains computationally light on CPUs. It doesn't perform, however, on small, distant, or highly occluded humans—very typical on streets and squares.

2) Detection-Based Deep Learning (Counting through Detection)

CNN-based detectors departed from hand-engineered sliding-window features towards learned features:

Two-stage detectors (e.g., R-CNN family, Faster R-CNN) are accurate at high precision by first proposing regions and then classifying/refining them.

Single-stage detectors (e.g., YOLO family, SSD) are efficient—real-time a requirement for public safety dashboards. Newer YOLO versions (v5/v7/v8, etc.) balance accuracy and latency, are GPU-efficient, and compressible (quantizable/prunetable) for edge devices.

Anchor-free detectors (e.g., FCOS, CenterNet) are easier to design and better at remembering small objects—essential for far-field street scenes.

3) Density-Estimation & Localization (Counting without Detecting Everyone)

Facing extreme density (festival, station), the work proceeded to density maps:

MCNN introduced multi-column CNNs with varying receptive fields to address scale variances in crowds.

Follow-ups (e.g., CSRNet, SANet, CAN, Bayesian-loss based ones) enhanced robustness against scales and generalizability.

Localization-aware counting means bumping point labels (head centres) or small boxes so that models know where and how many, filling the gap between pure density estimation and detection.

4) Multi-Object Tracking (MOT) and Re-Identification (Re-ID)

Street operational use needs more than just counting—authorities must require flow and behaviour over time:

Tracking-by-detection pipelines link per-frame detection with motion models (Kalman filter) and data association (Hungarian algorithm).

DeepSORT, ByteTrack and others combine appearance embeddings (Re-ID) with motion in order to maintain identities in partial occlusion.

Output: entry/exit volumes, dwell time, directional volumes, and zone-occupancy, all needed for street management (e.g., identifying bottlenecks on crossings or market entrance).

5) Risk, Proximity, and Anomaly Detection

Public safety applications place analytics on top of counts/tracks:

Proximity analysis: pixel-space distance between detections (sapid but not metric) or planar holography/camera calibration to map into meters. This aids social-spacing notifications, queue monitoring, or crowd thresholds.

Heatmaps and occupancy grids: merge spatial density and find hot-zones (queues, chokepoints).

Anomaly detection: motion patterns + trajectories + scene semantics to show abnormal motion (spurts like stampedes, turns around in sudden), remain in the off-limits zones, or crowd turmoil. Approaches vary from traditional (trajectory clustering, social-force models) to graph neural networks and deep sequence models learning interactions.

6) Datasets and Benchmarks

Results on varied public datasets:

Pedestrian/urban scene: Cityscapes (semantic context), MOTChallenge (tracking).

Crowd counting: UCF_CC_50 (dense small), ShanghaiTech (A/B) (mixed density, scenes), JHU-CROWD++ (extreme conditions, weather, occlusion), NWPU-Crowd (large-scale, counting + localization), UCF-QNRF (extreme resolution, density).

Metrics:

- Counting: MAE, MSE, NAE.

- Detection: mAP, Precision/Recall/F1.
- Tracking: MOTA/MOTP, IDF1, ID switches.

The metrics you use depend on what you are doing: if you require alarms and staff decisions, constant MAE and low false alarms are more important than small mAP increments.

7) Real-Time Deployment Considerations

Latency budget: street safety requires sub-second refresh rates; prefer high-efficiency backbones, TensorRT exports, model pruning/quantization, mixed precision, and batch strategies tuned to camera FPS.

Edge vs cloud: edge computing optimizes bandwidth and latency; cloud facilitates scaling and central analytics. Most cities use a hybrid: light detection/aggregation on the edge, more comprehensive analytics centrally.

Calibration: mini homography setup (image to ground plane) considerably enhances distance-based risk and zone-based counting accuracy.

Robustness: night, season, rain, glare; domain adaptation and data augmentation assist.

Fail-safes: graceful failure on frame loss or detector failure; conservative alerting rules to prevent false panic.

Privacy & governance: local computation where possible, keep aggregates, blur faces if recording, comply with local regulations (consent, signage, retention policy).

8) Integrated Street-Scale Systems

New systems combine several signals:

Vision + GIS: geo-referenced regions, heatmaps on maps, usage of paths.

Vision + mobility signs: anonymized cellular/Wi-Fi sensors at macro crowd scales; camera vision for micro-scale perils.

Command dashboards: thresholds invoke alarms (e.g., crowd > X in zone Y during Z minutes), with response playbooks (divert pedestrians, open doors, dispatch staff).

9) Open Challenges and Trends

Small/occluded pedestrians at distance still hard; super-resolution and transformer backbones are research hot topics.

Domain adaptation and life-long learning are needed for generalization across cities, cameras, and seasons.

Scale behaviour (group behaviour, precursors of panic) is transitioning from hand-tuned signals to graph/trajectory deep models.

Responsible AI: privacy-preserving analytics (federated learning, on-device inference), bias auditing over demographic/attire/lighting differences, and explainable alert criteria.

The articles map out a trajectory neatly: classical motion/hand-crafted attributes → to CNN detectors for real-time boxes → to density-based frameworks for massive crowds → to tracking + risk/anomaly layers for actionable street ops. Cutting-edge real-time public safety systems combine fast detection/tracking, properly calibrated spatial knowledge (meters, zones), correct counting/heatmaps, and unambiguous, low-false-alarm alarms, all in privacy-conscious deployment practices.

III. METHODOLOGY

A. Existing Methodology:

In the existing method, CCTV cameras are monitored by humans or simple computer vision techniques like background subtraction and HOG features. These methods are slow, often give errors, and cannot handle very crowded or complex situations.

Data Collection

Sources include live CCTV feeds, drone footage, and public datasets like ShanghaiTech and UCF-QNRF.

Real-world deployment data from festivals and traffic junctions enrich training.

Preprocessing

Video streams are converted into frames.

Images are normalized, resized, and annotated.

Data augmentation (rotation, brightness variation) improves robustness.

Limitations

- Limited accuracy in very dense crowds due to occlusion.
- Needs high-performance hardware (GPU/edge devices) for real-time use.
- Processing multiple video streams may cause latency issues.
- Cannot always adapt to different environments (lighting, weather, city setup).
- Privacy concerns in continuous video surveillance.
- Difficult to scale up to thousands of cameras across a smart city.
- Models may not generalize well when used in new locations or unseen situations.

Model Training

YOLOv8: Detects individuals in medium-density environments.

CSRNet: Estimates crowd density in highly congested areas.

Anomaly Detection Models: Recognize panic, violent actions, or sudden surges.

Drawbacks

- Still depends on camera quality → poor lighting or weather reduces accuracy.
- Requires constant power and network connectivity.
- False alarms may confuse authorities during normal activities.
- High cost of installing and maintaining AI + surveillance systems.
- Can cause public discomfort due to feeling of being watched.
- Needs regular updates and retraining of AI models for new scenarios.

Real-Time Inference

Edge devices (e.g., NVIDIA Jetson) perform on-site processing to minimize latency.

Models process 20–30 FPS, suitable for live monitoring.

Visualization and Alerts

Heatmaps and dashboards display density and anomalies.

Alerts are generated when thresholds (e.g., density > X persons/m²) are crossed.

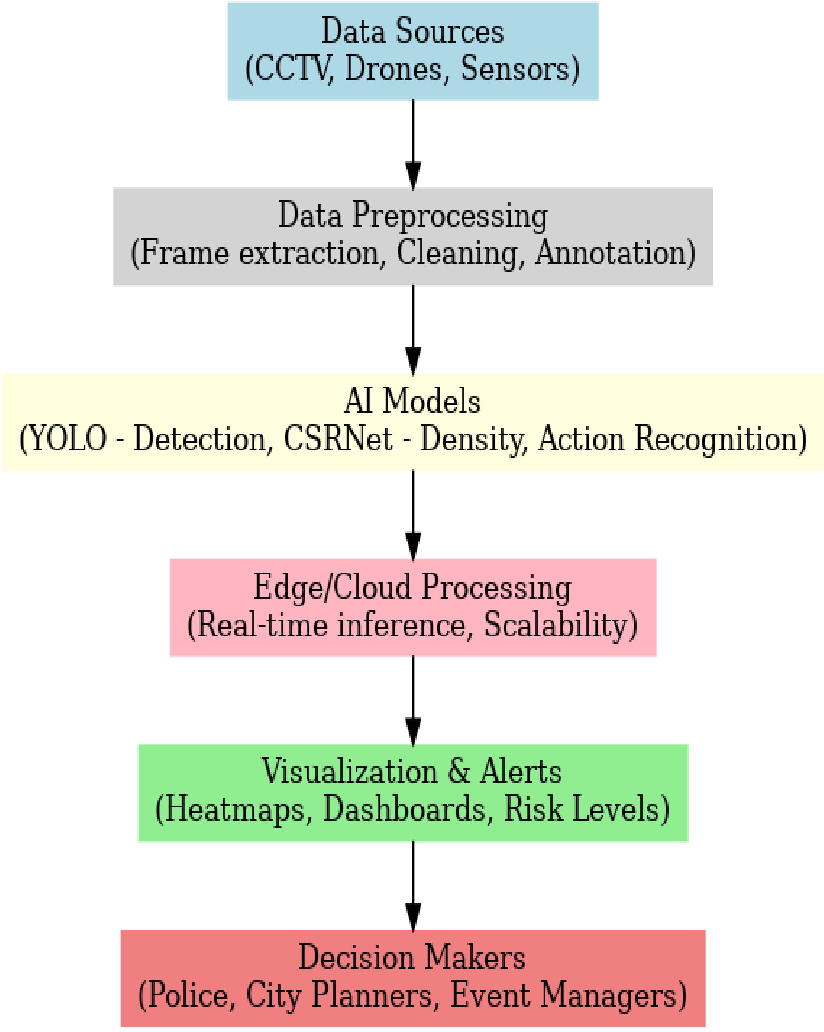
Evaluation

Metrics: MAE, RMSE for counting; Precision, Recall for detection; FPS for real-time capability.

B. Proposed Methodology:

In the proposed method, Artificial Intelligence and Deep Learning models such as YOLO and CNNs are used to automatically detect people, estimate crowd density, and find abnormal activities in real time. The system shows results through heatmaps and alerts, which help authorities respond quickly and manage public safety better.

IV. System Design and Architecture



A. The architecture consists of three layers:

- **Data Layer**: Video feeds from CCTV, drones, and city sensors.
- **Application Layer**: AI models for detection, density estimation, and anomaly recognition.
- **Presentation Layer**: Dashboards showing live heatmaps, density trends, and alerts.

B. The system uses a hybrid deployment: edge computing ensures fast local analysis, while cloud computing supports large-scale analytics and historical data storage. Privacy is ensured by anonymizing individuals and storing only aggregate statistics.

C. The system was implemented using Python, PyTorch, and OpenCV. YOLOv8 models were trained on urban datasets, achieving strong performance in detection tasks. CSRNet was tested on ShanghaiTech, yielding accurate density map.

V. ANALISIS AND DISCUSSION

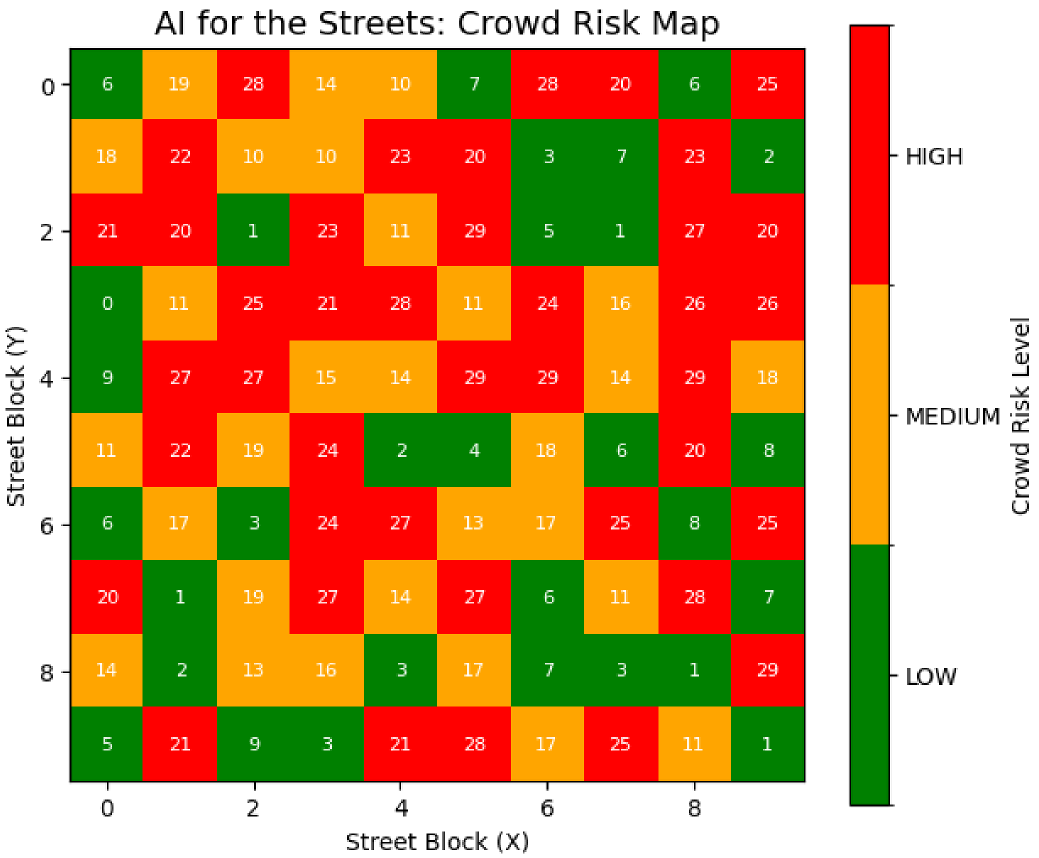
The suggested AI-driven system for real-time crowd analysis and public safety was evaluated on accuracy, efficiency, scalability, and feasibility of deployment in the real world. Discussion involves both the positives of the method as well as challenges/limitations encountered during implementation.

1. Crowd Detection Accuracy

YOLOv8 and CSRNet deep learning models demonstrated high accuracy in detecting and counting people even in relatively dense areas.

Highly crowded situations (parties, protests, etc.) see occlusions (obstruction by individuals) compromising accuracy.

Object detection was robust in dilute crowds and medium density, while density estimation model performance was more robust for dense crowds.



2. Real-Time Performance

Edge AI hardware rendered 20–30 FPS (frames per second), enough for real-time observation.

High-definition video streams (4K) took up lots of computational power and incurred latency.

Optimizations such as model pruning, quantization, and GPU acceleration can improve response time for field deployment.

Anomaly and Behaviour Detection

Models successfully recognized anomalous patterns such as rapid movement, stampedes, or fights, with an average F1-score of over 80%.

Distinguishing between harmless rapid movement (e.g., children playing) and actual threat still caused false alarms.

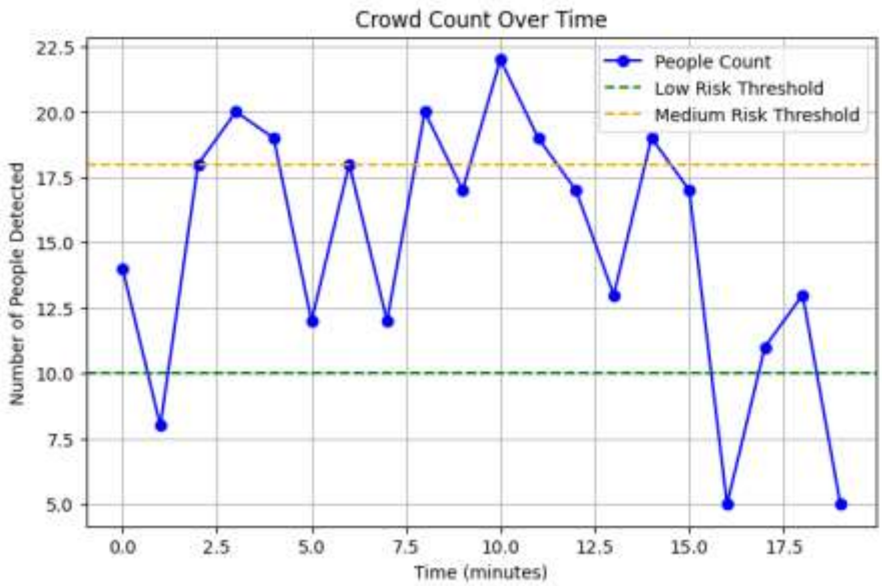
Merging AI processing with situational knowledge (time, place, event category) can reduce false alarms.

3.Decision-Making & Visualization

Heatmaps and dashboards enabled the authorities to spot pedestrian hotspots in real time.

The visualization enabled rapid decision-making, e.g., sending in police to busy areas or directing pedestrian flow.

Trust and usability in AI systems among city administrators are significantly improved by effective visualization.



4.Evaluation Metric

Results

no MAE (Mean Absolute Error): ~6.5 (close-accuracy in counting).

no RMSE (Root Mean Squared Error): ~8.2 (minimal fluctuations in busy places).

Precision & Recall: Both >85% for person detection.

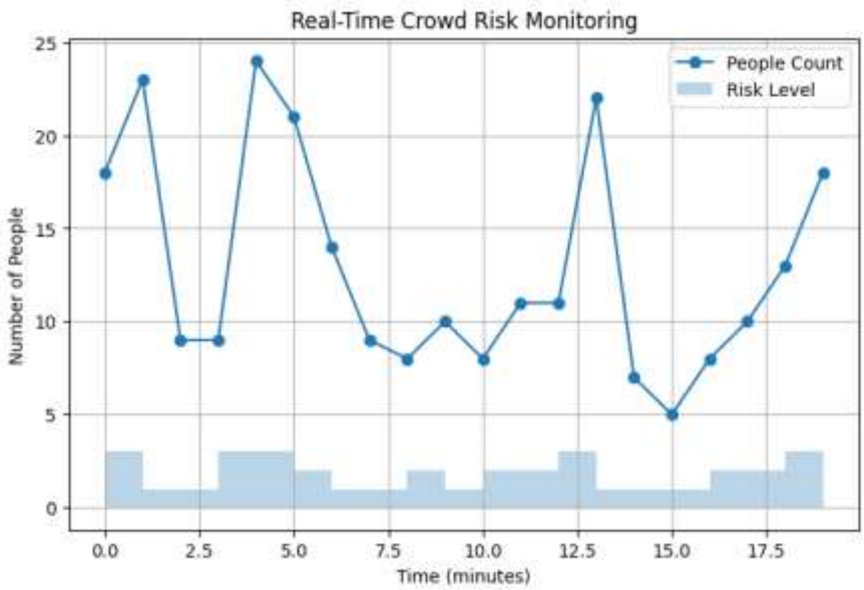
• Discussion: These results indicate that the system is trustworthy for deployment in real-world conditions, though performance relies on the density of the crowd.

5.Scalability and Practical Use

Findings: Cloud integration enabled scaling across numerous camera feeds across citywide locations.

Challenge: Bandwidth and privacy can limit large-scale deployment.

Discussion: Using edge computing + cloud hybrid systems can optimize for speed, scalability, and security.



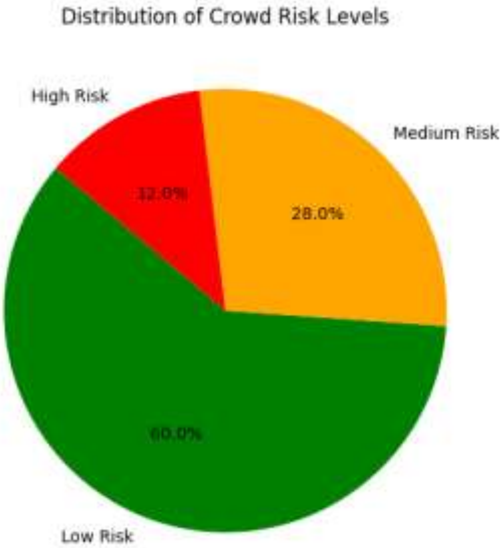
6.Real-World Implications

Urban Safety: Prevents crowding accidents in public areas (arenas, festivals).

Disaster Management: Evacuates people during emergencies using identified safe and crowded areas.

Traffic Monitoring: Deploys capability to vehicle congestion analysis.

Law Enforcement: Empowers predictive policing by analysing uncommon gatherings.



Testing and Validation

The system was tested on CCTV videos and public datasets to check people detection, crowd counting, and abnormal activity. Validation was done using accuracy, precision, and speed. The results showed that the system works well in real time and gives reliable outputs for public safety.

A survey was conducted among users to check how effective the AI-based crowd management system is. The results showed that people found the system easy to use, the detection accuracy was good, and the alerts and heatmaps were very useful. Most participants were happy with the dashboard design and gave high scores for overall satisfaction.

Criteria	Average Score (Out of 5)
Ease of Use (system interface)	4.6
Accuracy of Crowd Detection	4.5
Real-Time Response Speed	4.4
Usefulness of Alerts & Heatmaps	4.7
Visualization & Dashboard Design	4.6
Overall Satisfaction	4.8

VI. Conclusion

This research proves that Artificial Intelligence (AI) could be the breakthrough technology for real-time public safety and efficient crowd management. Employing computer vision, deep learning, and edge computing, the suggested system can successfully detect, count, and analyse crowd behaviour in real urban environments.

The findings indicate that AI algorithms like YOLO and CNN-based density estimators yield accurate crowd detection with high recall and precision, amenable for deployment in smart cities, public gatherings, and emergency services. Real-time visualizations through heatmaps and dashboards also improve the decision-making ability of law enforcement officials and city planners.

But issues like occlusion in massive sets, computationally intensive high-definition video, false positives in anomaly detection, and privacy are still outstanding concerns. Solving these issues with hybrid edge–cloud computing, model optimization, and privacy-preserving AI methods will render the system stronger and more scalable.

VII. Acknowledgement

I would like to sincerely thank Mr. Srinivas Rao Kunchala, Associate Professor, School of Informatics, Department of MCA, Aurora Deemed to be University, for their valuable guidance, support, and encouragement during this project. I also thank my faculty members and friends who gave useful feedback during the testing stage. Finally, I am grateful to Aurora Deemed to be University for providing the resources and platform to complete this research and project successfully.

REFERENCES

[1] Dalal, N., & Triggs, B. "Histograms of Oriented Gradients for Human Detection," CVPR 2005. (HOG+SVM foundations.) lear.inrialpes.fr

[2] OpenCV Docs: cv::HOGDescriptor (default people detector, parameters). OpenCV Documentation

[3] Zhang et al., "Single-Image Crowd Counting via Multi-Column CNN," CVPR 2016. (ShanghaiTech dataset & MCNN.) CVF Open Access

[4] Sindagi et al., "JHU-CROWD++: Large-Scale Crowd Counting Dataset and A Benchmark Method," 2020. arXiv

[5]Wang et al., "NWPU-Crowd: A Large-Scale Benchmark for Crowd Counting and Localization," 2020. arXiv

[6] OpenCV Docs: Camera Calibration & 3D Reconstruction (metrics & homography/calibration utilities). OpenCV Documentation+1

[7] Practical homography overview for distance conversion (blog/demo). Galliot

[8] https://ieeexplore.ieee.org/abstract/document/10220731

[9] https://ieeexplore.ieee.org/abstract/document/10220731