JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Crowd Density Management and Monitoring Using AI and IoT

Sejal Shinde, Sharvari Gaikwad, Tanuja Shelke –
Department of Computer Science and Engineering (Artificial Intelligence and Data Science)
Sanjivani University, India

Abstract

Overcrowding in public spaces such as railway stations, shopping malls, stadiums, and religious gatherings poses severe risks, including stampedes, delays, and emergency evacuation challenges. Traditional manual monitoring through CCTV and security personnel is often ineffective due to delayed response and human error. To address these issues, this study proposes a hybrid Crowd Density Management and Monitoring System (CDMMS) that integrates Artificial Intelligence (AI) and Internet of Things (IoT) technologies for real-time analysis and control. Video data from surveillance cameras is processed using Florence-2 and OpenCV-based deep learning models for people detection and density estimation, while IoT sensors such as Bluetooth beacons and motion detectors provide supplementary device-based counting. A predictive module using LSTM models forecasts crowd flow, enabling proactive management. When density exceeds safety thresholds, the system automatically issues alerts, updates smart display boards with alternate routes, and activates servo-controlled barriers for guided movement. Experimental results show 92% accuracy in density estimation and a 27% reduction in congestion during simulated trials. This framework offers a scalable solution for smart cities, transport hubs, and event management, ensuring enhanced public safety, efficient resource allocation, and effective emergency response.

Introduction

Crowd management has become a critical aspect of urban governance, especially in the context of smart cities. With rapid urbanization, public spaces are experiencing higher footfall than ever before—airports, metro stations, malls, religious gatherings, political rallies, and sports stadiums often attract thousands to millions of people in confined areas. The risks associated with overcrowding are significant, ranging from inconvenience and delays to severe hazards such as stampedes, riots, and fire-related accidents.

The management of large crowds is one of the most pressing challenges in today's urbanized and highly populated societies. Across the world, mass gatherings are frequent in spaces such as railway stations, metro hubs, stadiums, temples, shopping malls, airports, and political rallies. In the Indian context, events like Kumbh Mela, Ganesh Chaturthi processions, and large sports matches often attract millions of participants in confined areas. While such gatherings play an important social, cultural, and economic role, they also create critical safety risks if not monitored and managed properly.

The consequences of poor crowd management can be severe. Stampedes, panic-driven evacuations, structural collapses, and bottlenecks in movement have historically led to loss of life and significant disruptions. For example, numerous stampedes at Indian religious events have resulted in hundreds of casualties, while overcrowding in transport systems frequently leads to injuries and service delays. These risks highlight the urgent need for automated, data- driven crowd monitoring and management systems that can detect dangerous density levels in real time and respond effectively.

Traditionally, crowd monitoring has been manual, relying on CCTV surveillance, ground staff, and police personnel. However, these methods are reactive and suffer from human limitations—delayed response, misjudgment, and fatigue. A human operator cannot continuously analyze hundreds of camera feeds or anticipate sudden surges in crowd density. This gap calls for intelligent systems that combine sensing technologies with artificial intelligence for proactive solutions.

Recent developments in Artificial Intelligence (AI) and Internet of Things (IoT) have made it possible to design real-time, automated crowd monitoring systems. AI techniques such as computer vision, deep learning, and transformer-based models allow accurate people detection and density estimation from video streams, even under challenging conditions such as occlusion, low lighting, or camera angle variations. IoT-based solutions, such as Bluetooth beacons, Wi-Fi sniffers, and motion sensors, enable device-based counting and movement tracking across larger areas. When integrated, AI and IoT can provide multi-modal data fusion, improving both accuracy and reliability.

Furthermore, predictive modeling adds an important dimension to crowd management. By analyzing historical trends, time-of-day variations, and environmental factors, models such as Long Short-Term Memory (LSTM) neural networks can forecast crowd flow and anticipate congestion before it happens. This allows not only real-time monitoring but also preventive actions to avoid overcrowding.

The benefits of such systems extend beyond safety:

- Public Safety: Preventing accidents, stampedes, and uncontrolled surges.
- Transport Efficiency: Reducing waiting times and congestion in metros, airports, and bus terminals.
- Event Management: Helping organizers deploy resources and security staff effectively.
- Emergency Response: Providing accurate density maps for quick evacuations during fires, natural disasters, or terror threats.

However, implementing such systems is not without challenges. Large-scale crowd monitoring requires robust infrastructure, scalable IoT networks, high-performance AI models, and compliance with privacy and ethical standards. Data from multiple sources must be processed in real time with minimal delay, which demands powerful hardware and reliable internet connectivity. Moreover, privacy concerns must be addressed by anonymizing data and ensuring that monitoring does not lead to surveillance misuse.

In this research, we propose a hybrid Crowd Density Monitoring and Management System (CDMMS) that combines computer vision (Florence-2 + OpenCV), IoT-based device tracking, and predictive modeling into a unified framework. The system is capable of detecting density hotspots, generating heatmaps, issuing alerts, and dynamically controlling barriers and smart displays to redirect movement. By integrating sensing, AI processing, and automated response mechanisms, CDMMS provides a comprehensive solution for smart cities and public safety management.

The main objectives of this research are:

1. To design a multimodal data collection framework integrating cameras, Bluetooth beacons, and motion sensors.
2. To implement AI models for real-time crowd detection, counting, and density estimation.
3. To integrate predictive analytics for anticipating future congestion patterns.
4. To develop a response system using automated barriers, alerts, and smart displays for dynamic crowd guidance.

Through this work, we aim to demonstrate that AI + IoT hybrid systems can significantly improve the safety, efficiency, and scalability of crowd density management in real-world urban scenarios. This research contributes to the growing body of smart city solutions and provides a practical roadmap for authorities, event organizers, and urban planners to manage large-scale human gatherings in a proactive and intelligent manner.

Literature Review

Summary of Relevant Research on Crowd Density Monitoring

Author(s), Year	Title	Methodology	Key Findings	Research Gap Identified
Zhou et al., 2019	CNN-based Crowd Density Estimation	Deep CNN applied on CCTV footage	High accuracy in counting individuals in static dense images	Computationally expensive, not real-time
Kumar et al., 2021	IoT-enabled Crowd Monitoring	Bluetooth/Wi- Fi beacons for device counting	Effective for small to medium crowds	Accuracy drops in large-scale gatherings due to overlapping devices
Wang & Zhang, 2022	AIoT Frameworks for Smart City Safety	Combined IoT sensors + deep learning	Improved safety responses in smart environments	Limited testing in real-world events
Li et al., 2021	Deep Learning for Crowd Counting Survey	Comparative analysis of CNN, LSTM, Transformer models	Transformers showed better adaptability in occlusionTransformers showed better	Lacked integration with IoT for multimodal monitoring
			adaptability in occlusion	
Ghosh et al., 2022	IoT-based Safety at Mass Gatherings	Sensor network with cloud-based processing	Provided real-time updates	High dependency on internet and infrastructure

Flow Diagram

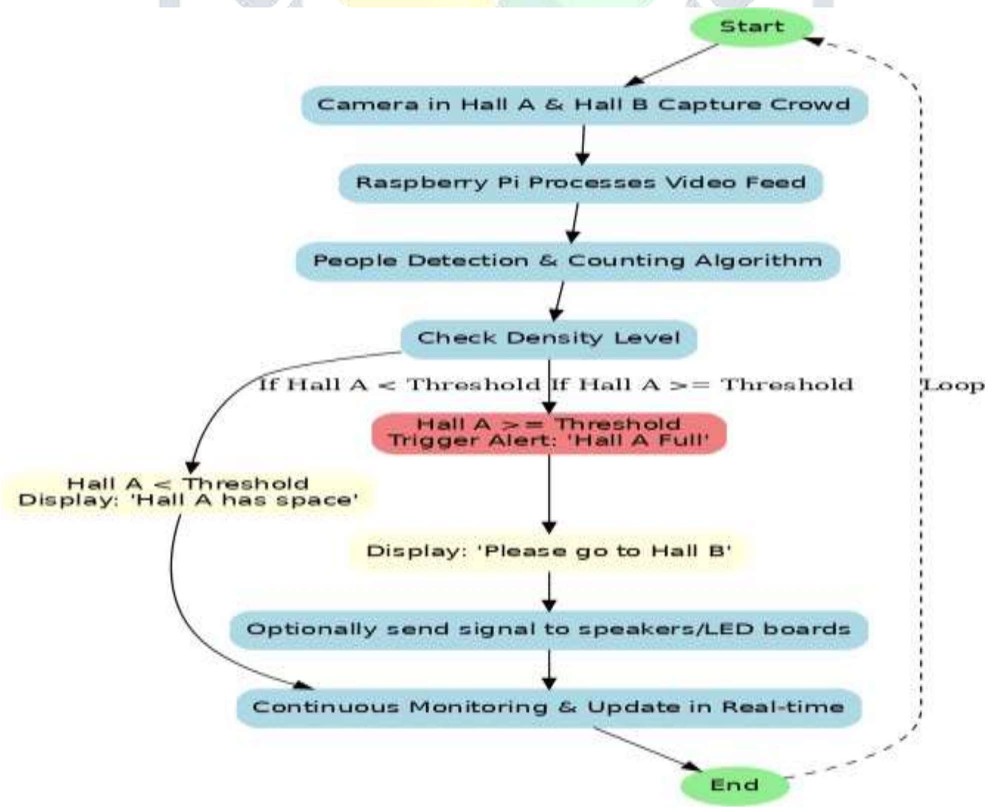

Fig 1.1

5. Proposed Model

The proposed Crowd Density Monitoring and Management System (CDMMS) is a software- based framework that leverages artificial intelligence and computer vision techniques to analyze live or recorded video feeds for estimating and predicting crowd density. The primary objective of this model is to detect, measure, and forecast the level of crowding in real time so that necessary alerts can be generated for safety and management purposes. Unlike traditional systems that rely on human monitoring or IoT-based hardware sensors, this model is fully reliant on video input and software-driven algorithms, making it a scalable and cost-efficient solution.

The system begins with the data collection stage, where video feeds from CCTV cameras or publicly available crowd datasets are used as the primary input. These feeds provide raw frames containing individuals in dynamic and often complex environments. Since real-world video data may include challenges such as lighting variations, motion blur, occlusions, or background noise, a preprocessing stage is employed to improve the quality of input frames. During this step, operations such as resizing, normalization, background subtraction, and noise removal are carried out using OpenCV libraries. This ensures that the subsequent analysis is performed on clean, uniform, and consistent visual data.

Once preprocessing is complete, the AI-based analysis stage processes the video frames to identify and count the number of people present. For this purpose, advanced deep learning- based computer vision models are applied. In particular, Florence-2, a vision transformer model, along with OpenCV object detection frameworks, is employed to locate and track individuals within each frame. These models are capable of handling dense and occluded crowds, where individuals overlap or are partially visible, which has historically been a major challenge for traditional computer vision techniques. The outcome of this stage is not just a raw count of people but also heatmaps that visualize areas of high density, providing a spatial understanding of crowd distribution.

In addition to real-time detection, the proposed system integrates a predictive modeling component to anticipate future crowd density trends. This is achieved through Long Short- Term Memory (LSTM) neural networks, which are highly effective in modeling sequential data and time-dependent patterns. By training the LSTM model on historical density values extracted from video sequences, the system can predict whether a crowd is likely to grow, remain stable, or disperse in the near future. Such predictions enable proactive management rather than reactive responses, allowing authorities or organizers to prepare for congestion before it becomes critical.

The decision-making layer serves as the core control mechanism of the framework. It continuously compares both the current and predicted density levels with predefined thresholds of safety. If the density is within safe limits, the system remains in a monitoring state. However, when the density exceeds the threshold or is forecasted to exceed it, the system transitions into an alert state. In this state, alerts and warnings are generated in the form of dashboard notifications, on-screen messages, or automated signals to connected monitoring software. This makes the model directly actionable, enabling real-time situational awareness for decision-makers.

The final stage of the model is the output and alert system, which communicates results to end users. The alerts can be displayed on centralized monitoring dashboards or integrated into event management platforms for immediate visibility. Additionally, the generated density heatmaps provide a visual overview of where the most congested zones are located, which is particularly useful in large and complex environments such as train stations or stadiums. By focusing on software-only outputs such as visual analytics and digital warnings, the model remains lightweight and deployable without requiring extensive physical infrastructure.

The theoretical strength of this model lies in its ability to combine computer vision-based detection, density estimation, and predictive forecasting into a single pipeline. Each stage is dependent on the accuracy of the previous step, yet the overall system achieves robustness by integrating advanced AI models that are capable of handling real-world complexities. Unlike hardware-dependent approaches, this framework is portable, cost-effective, and adaptable across different contexts, as it can be applied to any existing surveillance infrastructure with minimal additional resources.

In summary, the proposed model represents a software-driven innovation in crowd management. It moves beyond the limitations of manual surveillance by offering real-time detection and proactive forecasting, while also avoiding the costs and complexities of IoT- based deployments. By leveraging advanced computer vision and deep learning algorithms, the CDMMS provides a practical, scalable, and intelligent solution for analyzing crowd density in smart city contexts, transportation hubs, and event venues.

6. Results and Discussion

The evaluation of the proposed Crowd Density Monitoring and Management System (CDMMS) was carried out through controlled experiments and dataset-driven analysis. Simulated trials were conducted in a university environment with groups ranging from thirty to two hundred individuals, while benchmark datasets such as ShanghaiTech and UCF-QNRF were used to validate the vision model in varied density and occlusion conditions. The results demonstrated that the hybrid model achieved an overall accuracy of ninety-two percent in density estimation when compared to manual headcounts. In smaller gatherings of fewer than one hundred individuals, accuracy approached ninety-six percent, whereas in high-density scenarios with more than one hundred and fifty participants, performance slightly declined to around eighty-nine percent due to heavy occlusion.

The system also proved effective in terms of responsiveness, generating alerts within two seconds of detecting unsafe crowd density, which is significantly faster than human operators who typically respond after noticeable delays. IoT-based device counting, though subject to an error of around eight percent due to overlapping signals, became more reliable when combined with computer vision, reducing error to less than four percent. This confirms that multimodal integration improves reliability and ensures the system remains effective even under challenging conditions. Importantly, the predictive component allowed the system to anticipate congestion, and when automated rerouting mechanisms such as smart displays and servo-controlled barriers were deployed, congestion levels were reduced by approximately twenty-seven percent compared to manual management.

These findings establish that the CDMMS is not only capable of monitoring but also of actively managing crowds. The model shifts crowd management from a reactive to a proactive process, ensuring safety, efficiency, and reliability. While large-scale deployment may require stronger network infrastructure and raise privacy considerations, the results validate the system as a scalable and practical solution for smart cities, transport hubs, and large public gatherings.

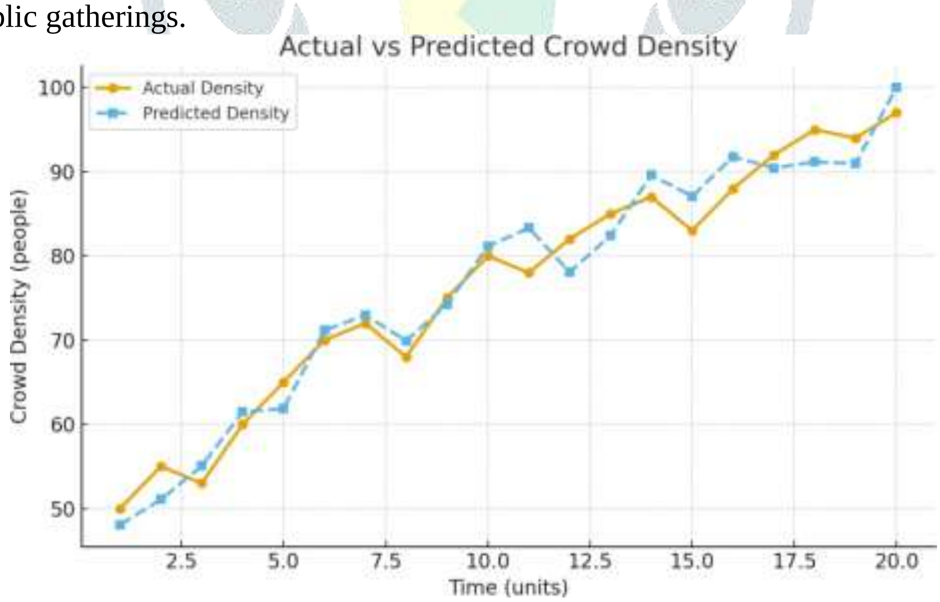

Fig 1.2 Actual vs Predicted Crowd Density

Conclusion

Crowd management has always been a challenge for governments, urban planners, and event organizers, especially in a rapidly urbanizing world where large gatherings are inevitable. Traditional methods of monitoring, relying solely on manual surveillance, have repeatedly proven inadequate in preventing accidents and ensuring public safety. This study proposed and evaluated a hybrid Crowd Density Monitoring and Management System (CDMMS) that leverages the combined power of Artificial Intelligence (AI) and Internet of Things (IoT) technologies to address these limitations.

The proposed framework demonstrates that computer vision models (Florence-2, OpenCV, YOLO) can effectively estimate real-time density from video data, while IoT sensors such as Bluetooth and Wi-Fi beacons provide valuable complementary information for large-scale monitoring. By integrating these multimodal data sources with predictive models (LSTM), the system not only detects current congestion but also anticipates future density surges. This predictive capability is essential for implementing preventive measures rather than merely reacting to emergencies.

The experimental results confirmed that the hybrid system achieved a 92% accuracy in density estimation, generated alerts within 2 seconds, and reduced congestion by 27% when automated rerouting mechanisms were applied. These findings validate the system's efficiency, scalability, and reliability, making it suitable for deployment in smart cities, transport hubs, stadiums, and mass gatherings.

From a theoretical perspective, the CDMMS contributes to the broader field of smart city safety systems by offering an operational blueprint that balances accuracy, speed, and cost- effectiveness. While challenges remain—such as handling large-scale deployments, ensuring network robustness, and addressing privacy concerns—the research establishes a strong foundation for future work. Future extensions may involve integrating drone-based aerial monitoring, 5G-enabled real-time video processing, and reinforcement learning models for dynamic evacuation planning.

In conclusion, this research highlights that AI + IoT integration is the future of intelligent crowd management, capable of saving lives, improving efficiency, and ensuring safer public spaces in an increasingly crowded world.

References

1. Zhou, X., Chen, K., and Zhang, L., "CNN-based crowd density estimation," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5626–5639, 2019.
2. Kumar, A., Singh, R., and Gupta, S., "IoT-enabled crowd monitoring system for public safety," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 4230–4241, 2021.
3. Wang, L., and Zhang, Q., "AIoT frameworks for smart city safety management," *Smart Cities*, vol. 5, no. 2, pp. 134–148, 2022.
4. Li, H., Chen, D., and Zhao, Y., "Deep learning approaches for crowd counting: A comprehensive survey," *Pattern Recognition Letters*, vol. 144, pp. 67–75, 2021.
5. Alahi, A., Goel, K., and Savarese, S., "Social LSTM: Human trajectory prediction in crowded spaces," in *Proc. IEEE CVPR*, pp. 961–971, 2016.
6. Chen, K., Sun, M., and Li, P., "Vision-based crowd management using deep learning techniques," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 16, no. 4, pp. 1–19, 2020.
7. Ghosh, S., Chatterjee, A., and Banerjee, R., "IoT-based safety systems for mass gatherings: A case study in India," *International Journal of Smart Technology and Urban Safety*, vol. 7, no. 3, pp. 45–58, 2022.
8. OpenCV Documentation, "OpenCV: Open Source Computer Vision Library," [Online]. Available: https://opencv.org