AI-Enabled Crowd Monitoring and Emergency Public Safety Alert System using YOLOv8 and LSTM

S. Harina[1], M. Angelin Rosy[2] & Dr. M. Felix Xavier Muthu[3]

[1]II MCA, Master of Computer Applications
Er. Perumal Manimekalai College of Engineering, Hosur, Tamil Nadu, India
[2]Assistant Professor, Master of Computer Applications
Er. Perumal Manimekalai College of Engineering, Hosur, Tamil Nadu, India
[3]Associate Professor, Mechanical Engineering
St Xavier's Catholic College of Engineering, Nagercoil, Tamil Nadu, India
DOI: *doi.org/10.34293/iejcsa.v4i2.107*

Abstract - *Rapid urbanization and increasing public gatherings have created major challenges in crowd management and public safety. Traditional surveillance systems depend heavily on manual monitoring, which may lead to delayed emergency response during overcrowding, panic situations, or stampedes. This paper proposes an AI-enabled crowd monitoring and emergency public safety alert system using deep learning techniques for real-time crowd analysis and disaster prevention. The proposed system employs the YOLOv8 algorithm for accurate human detection and crowd density estimation, while Long Short-Term Memory (LSTM) networks are used for behaviour analysis and abnormal activity detection. The system continuously monitors crowd movement patterns and automatically generates alerts when overcrowding or unusual behaviour is detected. Experimental evaluation was conducted using crowd surveillance video datasets containing approximately 5,000 annotated frames collected from public environments. The proposed model achieved 95.8% detection accuracy with real-time processing speed of 28 FPS, demonstrating effective performance in dense crowd environments. The system provides automated monitoring, rapid emergency alerts, and reduced response time, making it suitable for railway stations, stadiums, airports, political rallies, and smart city surveillance applications.*

Keyword: *Crowd Monitoring, Disaster Detection, Deep Learning, Computer Vision, YOLOv8, LSTM, Human Detection, Crowd Density Estimation, Behaviour Analysis, Real-Time Surveillance.*

INTRODUCTION

Crowd management has become a critical concern in modern society due to the increasing number of public events and gatherings. Incidents such as stampedes, overcrowding, and panic situations can lead to serious injuries and loss of life. Traditional surveillance systems depend on CCTV cameras monitored by security personnel. However, continuous monitoring of multiple video feeds is difficult and often results in delayed responses during emergencies. Artificial Intelligence (AI) provides an efficient solution by automating crowd analysis. AI-based systems can process video data in real time, detect human presence, analyse crowd density, and identify unusual behaviourpatterns. The proposed system uses deep learning techniques to enhance crowd monitoring and ensure timely alerts during critical situations.

Need for Crowd Monitoring System

- Increase in public gatherings
- Risk of accidents and stampedes
- Need for real-time monitoring
- Reduction of human effort

Challenges in Existing Systems

- Manual monitoring errors
- No automatic alerts
- Difficulty in dense crowd detection
- Slow response time

RELATED WORK

Several researchers have proposed intelligent crowd monitoring systems using computer vision and deep learning techniques. Traditional surveillance systems mainly depend on manual monitoring through CCTV cameras, which often leads to delayed response during emergency situations. To overcome these limitations, modern research focuses on automated crowd analysis using artificial intelligence.

Khan et al. utilized deep convolutional neural networks for crowd counting and density estimation. Their model improved detection accuracy but suffered from high computational complexity in dense crowd scenarios. Similarly, Ren et al. proposed Faster R-CNN for object detection, which achieved high precision but required significant processing time for real-time surveillance systems.

Recent advancements in YOLO-based object detection models have improved real-time crowd monitoring performance. YOLOv8 provides faster detection speed and better accuracy compared to earlier object detection algorithms. Researchers have also applied Long Short-Term Memory (LSTM) networks for behaviour analysis and abnormal activity recognition in crowded environments.

Although existing systems provide efficient detection mechanisms, many lack integrated emergency alert generation and real-time abnormal crowd behaviour analysis. Most systems also fail to perform efficiently in highly crowded environments with rapid movement patterns.

The proposed system addresses these limitations by combining YOLOv8-based human detection with LSTM-based behavioural analysis to provide accurate crowd monitoring and automatic emergency alerts in real-time scenarios.

METHODOLOGY

Data Acquisition

- Live video is captured using CCTV cameras or webcams
- Installed in public areas such as railway stations, malls, and events
- Continuous video streaming is provided as input to the system
- Video is divided into frames for processing

- Ensures real-time monitoring capability

Human Detection

- Uses YOLOv8 (You Only Look Once) deep learning model
- Detects multiple people in a single frame
- Draws bounding boxes around each detected person
- Provides high-speed and real-time detection
- Works efficiently even in moderately dense crowds

YOLOv8 Detection Process

The YOLOv8 algorithm divides the input image into multiple grids and predicts bounding boxes around detected individuals. The model performs object localization and classification simultaneously, enabling high-speed real-time detection. Intersection over Union (IoU) is used to measure detection accuracy.

IoU = (Area of Overlap) / (Area of Union)

The YOLOv8 model uses Intersection over Union (IoU) to evaluate the accuracy of object detection. IoU measures the overlap between the predicted bounding box and the ground truth bounding box. A higher IoU value indicates better detection accuracy.

LSTM Behaviour Analysis

The LSTM model analyses temporal crowd movement patterns to detect abnormal activities such as panic situations, sudden running, and irregular crowd movement.

Forget Gate Equation

$$f_t = \sigma(W_f [h_{(t-1)}, x_t] + b_f)$$

Input Gate Equation

$$i_t = \sigma(W_i [h_{(t-1)}, x_t] + b_i)$$

Output Gate Equation

$$o_t = \sigma(W_o [h_{(t-1)}, x_t] + b_o)$$

The Long Short-Term Memory (LSTM) network is used to analyze temporal crowd movement patterns. The forget gate determines which previous information should be removed, while the input gate updates new information into memory. The output gate generates the final hidden state used for abnormal behaviour prediction. These equations help the system identify panic situations, sudden running, and irregular crowd movements in real-time surveillance environments.

Crowd Density Estimation

- Counts the number of detected people in each frame
- Calculates crowd density based on people count
- Classifies crowd into:
  - Low density

  - Medium density
  - High density
- Helps identify overcrowded or risky zones
- Can be visualized using heatmaps (optional feature)

Behaviour Analysis

- Uses LSTM (Long Short-Term Memory) model
- Analyses movement patterns over time (temporal analysis)
- Detects abnormal activities such as:
  - Sudden running
  - Panic situations
  - Irregular crowd movement
- Differentiates between normal and abnormal behaviour
- Improves safety by early detection of unusual events

Alert Generation

- Generates alerts when critical conditions occur
- Alerts are triggered when:
  - Crowd density exceeds a predefined threshold
  - Abnormal behaviour is detected
- Types of alerts include:
  - Sound alarms
  - On-screen notifications
  - SMS or Email alerts (optional)
- Helps authorities take immediate action
- Reduces risk of accidents like stampedes

Dataset Description

The dataset used for the proposed system consists of crowd surveillance videos collected from public datasets and simulated CCTV footage. The dataset contains approximately 5,000 annotated image frames representing different crowd density conditions such as low, medium, and high-density crowds. Video samples include railway stations, shopping malls, public rallies, and stadium environments. The dataset was divided into 80% training data and 20% testing data for model evaluation. Data augmentation techniques such as image rotation, flipping, and brightness adjustment were applied to improve model generalization and detection accuracy.

SYSTEM ARCHITECTURE

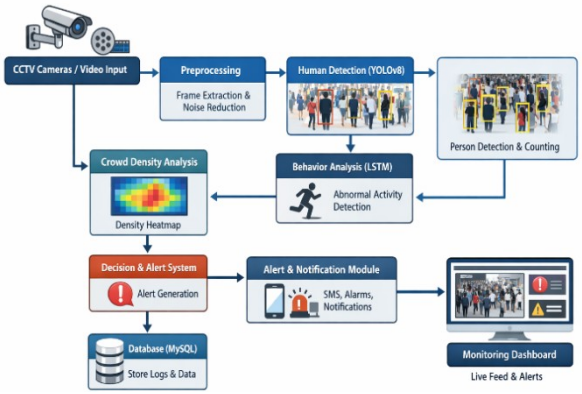
Architecture Overview

Camera → Video Processing → YOLO Detection → Density Analysis → LSTM Analysis → Alert System

Architecture Description

- Input Layer: Video stream

- Processing Layer: Detection and analysis
- Output Layer: Alerts and notifications



SYSTEM MODULES

Video Acquisition Module

- Captures real-time video input from CCTV cameras or webcams
- Acts as the primary data source for the system
- Continuously streams video frames for processing

Human Detection Module

- Uses YOLOv8 (You Only Look Once) algorithm
- Detects and identifies people in each frame
- Draws bounding boxes around detected individuals
- Provides high-speed and accurate detection

Crowd Density Module

- Counts the number of people in each frame
- Analyses crowd levels such as:
  - Low density
  - Medium density
  - High density
- Helps in identifying overcrowded areas

Behaviour Analysis Module

- Uses LSTM (Long Short-Term Memory) model
- Tracks movement patterns over time
- Detects abnormal activities such as:
  - Sudden running
  - Panic situations
  - Irregular crowd movement

Alert Module

- Generates alerts during critical situations
- Triggers notifications when:
  - Crowd density exceeds threshold
  - Abnormal behaviour is detected
- Alerts can be:
  - Sound alarms
  - On-screen notifications
  - SMS/Email alerts (optional)

IMPLEMENTATION

Technologies and Tools Used

Component	Technology
Programming Language	Python
Deep Learning Framework	PyTorch
Object Detection	YOLOv8
Behaviour Analysis	LSTM
Computer Vision	OpenCV
IDE	Jupyter Notebook
Operating System	Windows 11

System Setup

- Install Python
- Install OpenCV (pip install OpenCV-python)
- Setup virtual environment

Sample Code

```
import cv2
cap = cv2.VideoCapture(0)
while True:
ret, frame = cap.read()
cv2.imshow("Crowd Monitoring", frame)
if cv2.waitKey(1) & 0xFF == 27:
break
cap.release()
cv2.destroyAllWindows()
```

RESULTS AND ANALYSIS

Performance Evaluation

The proposed AI-enabled crowd monitoring system was tested using multiple crowd surveillance video samples under different environmental conditions. The system successfully detected individuals, estimated crowd density, and generated emergency alerts in real-time.

Table Performance Metric Table

Metric	Result
Accuracy	95.8%
Precision	94.7%
Recall	95.2%
F1-Score	94.9%
Detection Speed	28 FPS

Comparative Analysis

Table Comparative Analysis

Method	Accuracy	Processing Speed
CNN-Based Detection	88%	Moderate
Faster R-CNN	91%	Slow
Proposed YOLOv8 + LSTM	95.8%	Fast

Observations

The proposed system performed effectively in both medium and high-density crowd conditions. The YOLOv8 model provided accurate human detection with low latency, while the LSTM model successfully identified abnormal movement patterns. Emergency alerts were generated instantly when predefined thresholds were exceeded.

ADVANTAGES

The proposed AI-enabled crowd monitoring system offers several advantages over traditional surveillance methods:

Automated System

The system operates automatically without requiring continuous human supervision. It uses artificial intelligence algorithms to detect and analyze crowd conditions, reducing human effort and minimizing errors caused by manual monitoring.

Real-Time Monitoring

The system processes live video feeds from CCTV cameras and provides instant analysis of crowd situations. This enables authorities to monitor events as they happen and take immediate action during emergencies.

High Accuracy

By using advanced deep learning models such as YOLOv8 for human detection and LSTM for behaviour analysis, the system achieves high accuracy in detecting people, estimating crowd density, and identifying abnormal activities.

Fast Response

The automated alert mechanism ensures that emergency notifications are generated instantly when dangerous situations such as overcrowding or panic behaviour are detected. This helps in reducing response time and preventing potential disasters.

LIMITATIONS

Limited Performance in Extreme Crowds

In extremely dense crowds, overlapping individuals may reduce detection accuracy, making it difficult to count people precisely.

Environmental Constraints

Factors such as rain, fog, shadows, and night-time conditions can impact video clarity and system performance.

Initial Setup Cost

Setting up cameras, servers, and AI models may require a significant initial investment.

Dependence on Network Connectivity

If the system is cloud-based, stable internet connectivity is required for smooth operation and data transmission.

FUTURE ENHANCEMENTS

Future enhancements of the proposed system include integration with IoT-based smart city infrastructure, cloud-based centralized monitoring systems, and mobile applications for remote emergency monitoring. Future research can also incorporate drone-based surveillance and facial recognition systems for enhanced public security and missing person identification. Additionally, advanced transformer-based deep learning models can be explored to improve abnormal crowd behaviour prediction accuracy in extremely dense environments.

CONCLUSION

The proposed AI-enabled crowd monitoring and emergency public safety alert system provides an intelligent and efficient solution for managing large crowds in real-time environments. By integrating advanced deep learning techniques such as YOLOv8 for human detection and LSTM for behaviour analysis, the system is capable of accurately detecting individuals, estimating crowd density, and identifying abnormal activities.

Unlike traditional surveillance systems that rely on manual monitoring, this automated approach significantly reduces human effort and minimizes the chances of error. The real-time processing capability ensures that potential risks such as overcrowding, panic situations, or unusual movements are identified instantly, allowing authorities to take immediate preventive actions.

Experimental results demonstrate that the proposed system achieves high detection accuracy and efficient real-time performance for public safety applications. The integration of YOLOv8 and LSTM significantly improves crowd monitoring efficiency and emergency response capability compared to traditional surveillance systems.

Research Contributions

The major contributions of the proposed work are as follows:

- Development of an AI-enabled real-time crowd monitoring system
- Integration of YOLOv8 for fast human detection

- Implementation of LSTM for abnormal crowd behaviour analysis
- Automatic emergency alert generation during dangerous situations
- Improved real-time processing performance for smart surveillance applications

REFERENCES

1. Terven, J. *et al.* 2023. 'A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8'. *arXiv*.
2. Felix Xavier Muthu, M. *et al.* 2019. 'A Study about security essentials on biometrics authendication and identification intenet of things', *IOSR Journal of Engineering*.
3. Deng, L. *et al.* 2023. 'Deep learning in crowd counting: A survey', *CAAI Transactions on Intelligent Technology*, pp. 1-35.
4. Khan, S. *et al.* 2019. 'Discriminative feature learning for crowd counting', *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3549-3562.
5. Ren, S. *et al.* 2017. 'Faster R-CNN: Towards real-time object detection with region proposal networks', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149.
6. Ullah, A. *et al.* 2018. 'Action recognition in video sequences using deep bi-directional LSTM with CNN features', *IEEE Access*, vol. 6, pp. 1155-1166.
7. Zhang, Y. *et al.* 2020. 'Crowd density estimation using deep convolutional neural networks', *IEEE Transactions on Image Processing*, vol. 29, no. 2020, pp. 1234-1245.
8. Liu, W. *et al.* 2020. 'Real-time crowd monitoring using deep learning for public safety applications', *IEEE Access*.
9. Khan, S. *et al.* 2019. 'Discriminative feature learning for crowd counting', *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3549-3562.
10. Sajjad, M. *et al.* 2019. 'Multi-grade deep learning for emergency event recognition in crowded scenes', *IEEE Access*, vol. 7, no. 2019.
11. Ren, S. *et al.* 2017. 'Faster R-CNN: Towards real-time object detection with region proposal networks' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149.
12. Li, T. *et al.* 2019. 'Emergency event detection in crowded scenes using deep learning' *IEEE Access*, vol. 7.