Proactive Crowd Management System for Public Safety

AI-Powered Real-time Crowd Analysis and Congestion Forecasting

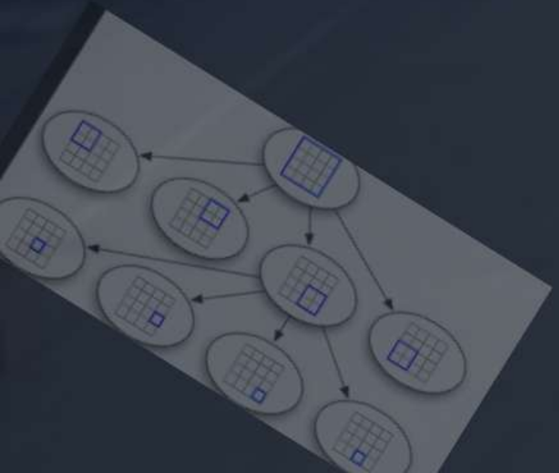
By:

Rachana Dixit (123B5F138)

Sohan Patil (123B5F139)

Guided By:

Dr. Harsha Bhute



Aim

To transform crowd management from a passive, reactive task into an intelligent, proactive process.
To accurately quantify crowds and forecast potential congestion to enable timely, proactive intervention.

The Limitations of Reactive Crowd Monitoring

Traditional public safety relies heavily on manual CCTV surveillance, which is inherently limited in dense, dynamic environments. This reactive approach often leads to interventions only after an incident has begun.

Manual & Reactive

Interventions occur only after critical thresholds or incidents, increasing risk of stampedes and injuries.

Prone to Human Error

Difficulty in accurately assessing density and tracking movement in highly complex and crowded scenarios.

Lack of Prediction

No capability to forecast emerging congestion hotspots or anticipate future public safety risks.

2025 Bengaluru crowd crush

Date	4 June 2025
Time	3:30 pm – 5:30 pm (IST)
Location	M. Chinnaswamy Stadium, Bengaluru, India
Type	Crowd crush
Cause	Overcrowding during celebrations
Deaths	11
Non-fatal injuries	56

Date	29 January 2025
Location	Prayagraj, Uttar Pradesh, India
Coordinates	25.431388°N 81.887971°E
Also known as	2025 Prayag Maha Kumbh Mela Crowd collapses
Type	Crowd collapses and Stampede
Cause	Broken barrier
Deaths	• 37 (official) • 37-82 (Sources disputed)
Non-fatal injuries	• 60 (official) • 90-200 (sources)

Motivation & SDG Goals

The Problem: Traditional reliance on manual CCTV surveillance is fundamentally reactive and susceptible to human error in dense scenes. It fails to prevent dangerous situations before they occur.

SDG 11 Focus: Our work is motivated by Sustainable Development Goal 11, which aims to build cities that are "inclusive, safe, resilient and sustainable".

Need for Intelligent Management

- Rapid urbanization increases the frequency and complexity of large crowd events.
- Intelligent public space management is crucial for large-scale event success and routine urban flow.

Technological Opportunity

- Transition from reactive observation to preventive safety measures is key.
- We leverage advanced computer vision and AI for automated crowd intelligence.

Objectives

1.REAL-TIME CROWD PERCEPTION

Implement efficient person detection using YOLOv11 on aerial imagery to ensure a reliable, low-latency data input stream.
Goal: Achieve the necessary competing goals of processing scenes in real-time and achieving accuracy during high crowding.

2. SPATIAL DENSITY ANALYSIS

Develop a Quadtree-based hotspot identification framework for localized congestion.
Goal: Efficiently and accurately pinpoint and identify current pressure points by allocating processing resources based on density.

3. PREDICTIVE FORECASTING

Create a GNN-based congestion prediction system using temporal Crowd Mobility Graphs (CMGraphs).
Goal: Forecast crowd movement and predict when and where crowding will develop, enabling crucial proactive intervention.

4. INTEGRATED ALERT SYSTEM

Design a proactive warning mechanism that combines real-time density alerts with predictive congestion insights.
Goal: Empower authorities with visual and textual alerts to prevent dangerous situations, transforming management from a reactive task into an intelligent, proactive process.

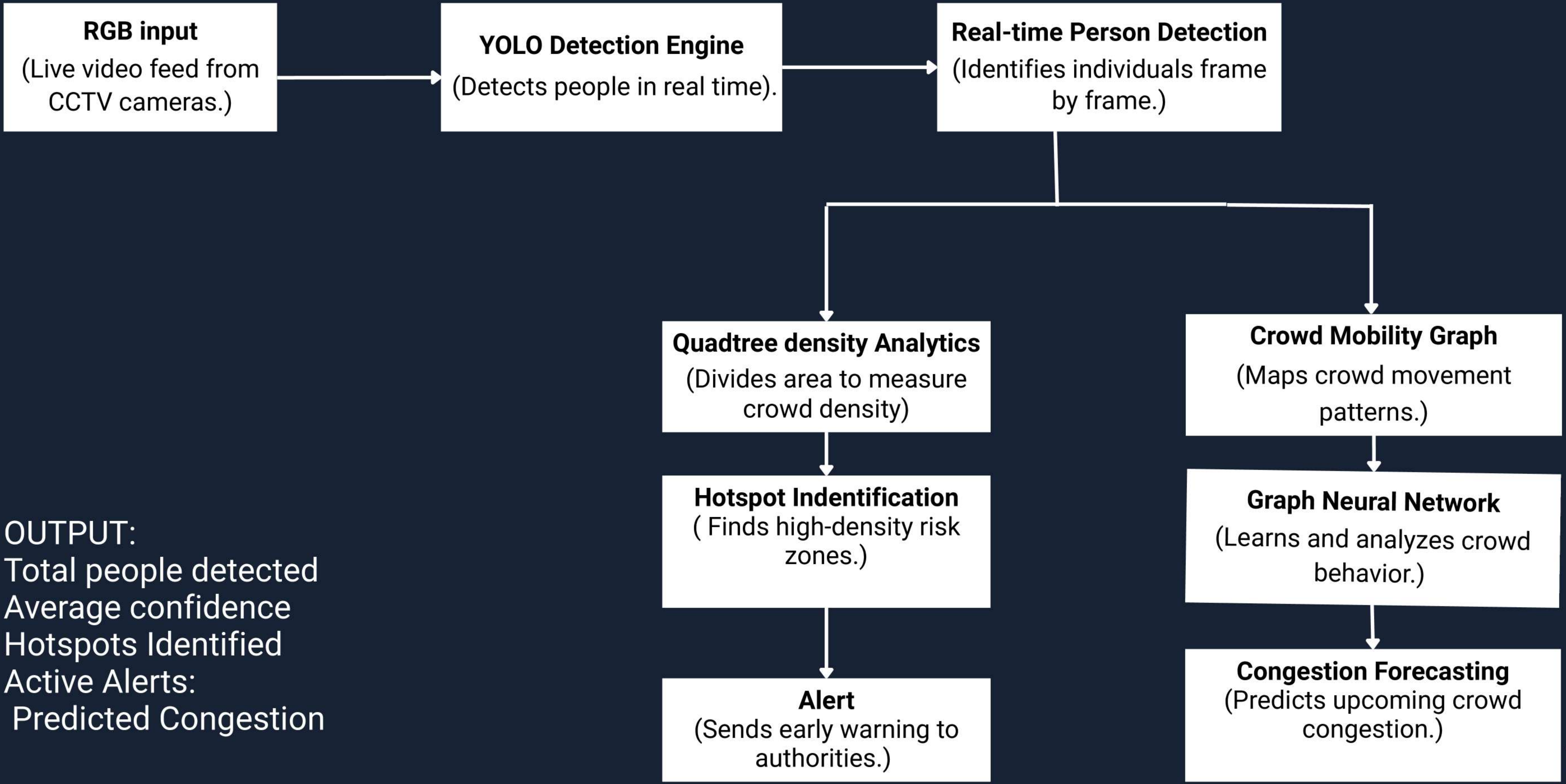
Literature Review

Author(s)	Methodology Focus	Findings / Key Contribution	Research Gap(s)
Zhuoxuan Peng	Single Domain Generalization (SDG) for crowd counting using MPCount.	Significantly improves counting accuracy, even in environments the model hasn't seen before, effectively solving the "domain shift" problem.	Focused purely on getting the count right; it did not offer a way to proactively forecast what the crowd would do next.
Yasiru Ranasinghe	Conditional Diffusion Models (CrowdDiff) with a regression branch.	Solves tough issues with background noise and density loss in crowd maps by using a narrow Gaussian kernel and generating multiple density maps.	Lacked integration with a broader, real-time, predictive framework and the continual generative synthesis of a dynamical density map.
Zhengyi Liu	Fused information from RGB and thermal cameras using a Transformer-based network.	Set a new standard for accuracy in poor lighting and challenging conditions by leveraging complementary multi-modal data.	Their system only works if both camera types are available, and it lacked any proactive forecasting components.
Yong-Chao	A weakly-supervised approach using a Parent-Child Network (PC-Net).	Achieved high accuracy while training only with image-level crowd counts (no precise location annotations), saving time and effort on manual data annotation.	The model is focused purely on density estimation and does not include the real-time tracking or predictive analysis needed for a full solution.
L.J.C. Valencia	Real-time crowd counting and tracking using the lightweight model Tiny-YOLOv4 and DeepSORT.	Demonstrated that a lightweight model is highly useful for a viable, real-time application in low-to-medium density crowds.	It is a reactive system—it only provides the current count and struggles significantly with high-density crowds.

Literature Review

Chengxin Liu	Developed the Point-query Transformer (PET) using a Quadtree structure.	Offered an elegant and smart way to dynamically process crowd regions based on density, making the model highly efficient.	The paper focused only on the core counting problem and did not show how to integrate it into a larger, comprehensive safety framework.
Yu Zhang	A novel anchor-free detector that dynamically learns the head-body ratio to focus on human heads.	This head-centric approach is highly effective for improving tracking accuracy and managing heavy occlusion in "extremely crowded scenes".	The work is limited to improving detection and tracking accuracy and does not offer a way to use that data for proactive forecasting.
Vivian W. H. Wong and Kincho H. Law	Conceptual framework using Crowd Mobility Graphs (CMGraphs) by fusing CCTV data and spatial floor plans.	Demonstrated that integrating visual and spatial data enables effective automation and proactive monitoring.	The framework remained conceptual and lacked an operationalized, non-concrete method for the real-time detection and counting that would serve as the framework's data input.
A comprehensive survey...	*Survey/Review Paper*	Provides a broad overview of methodologies and challenges in crowd density estimation and counting.	Not an original methodology paper; provides context rather than a specific technical gap.
Deep learning in crowd counting: A survey	*Survey/Review Paper*	Summarizes recent advancements and directions for deep learning models in crowd counting.	Not an original methodology paper; provides context rather than a specific technical gap.

Methodology

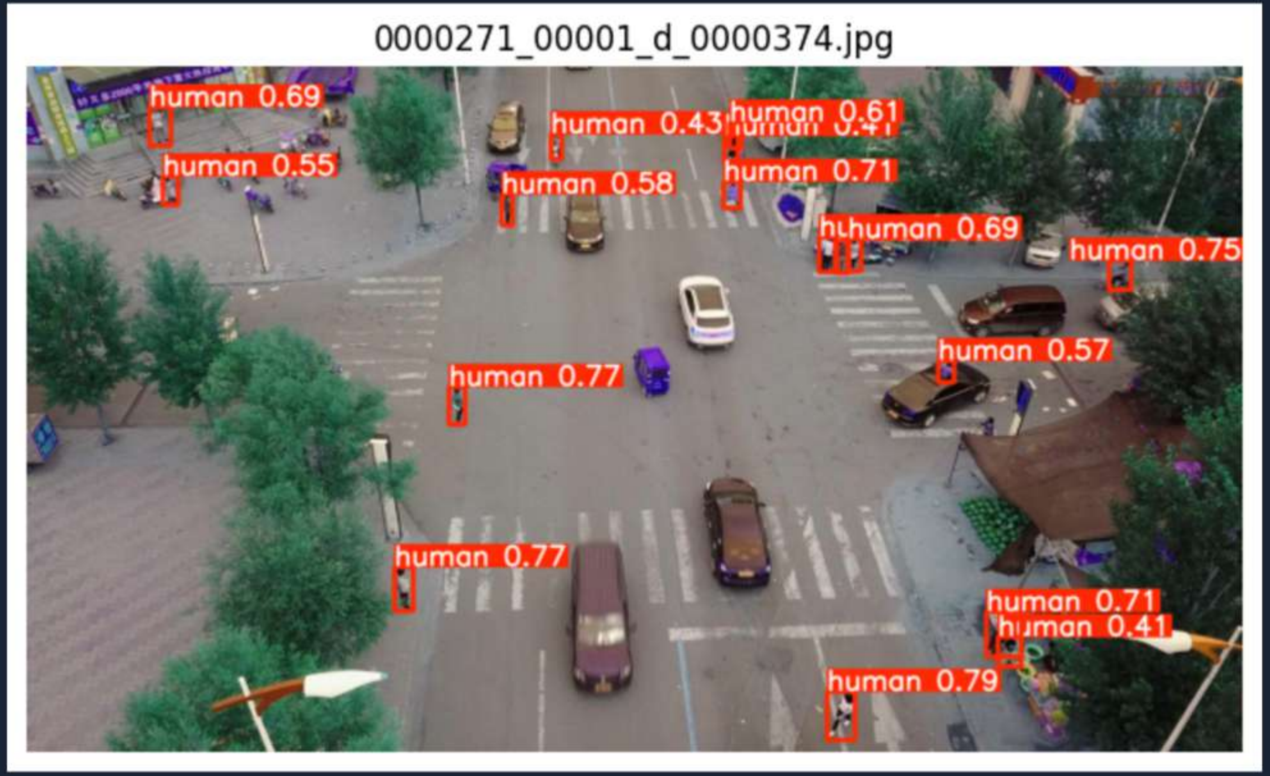


OUTPUT:
Total people detected
Average confidence
Hotspots Identified
Active Alerts:
Predicted Congestion

YOLOv11 Implementation

How YOLO Works:

- Single Pass Processing: Simultaneously predicts bounding boxes and class probabilities in one forward pass
- Grid-based Detection: Divides image into cells, each responsible for detecting objects within its area
- Anchor-free Design: Directly predicts object centers and sizes without predefined anchor boxes
- Multi-scale Feature Extraction: Uses CSPDarknet backbone to detect objects at different scales and resolutions



YOLO Implementation

What We Actually Built:

- Model: YOLOv11s fine-tuned on VisDrone2019 dataset
- Training Data: 6,471 training images + 548 validation images
- Performance: 45 FPS with 81% F1-score accuracy
- Detection Focus: Pedestrian class from aerial perspectives

Why we chose YOLOv11

- Our project deals with real-time proactive crowd monitoring — requiring speed + precision, not just accuracy.
- YOLOv4's heavy model could lag during live video processing or high frame-rate streams.
- YOLOv11 sustains >40 FPS with strong mAP (0.791) even in dense, aerial human scenes.
- Fine-tuning on VisDrone2019 dataset makes it ideal for top-down pedestrian detection, aligning perfectly with drone or CCTV surveillance angles.
- Maintains consistent detection under crowd density, lighting, and movement variations.

Technical Achievements:

- Successfully trained model for 45 epochs
- Achieved 0.791 mAP@0.5 detection accuracy
- Real-time processing capability maintained
- Robust performance across varying crowd densities

Validation Results:

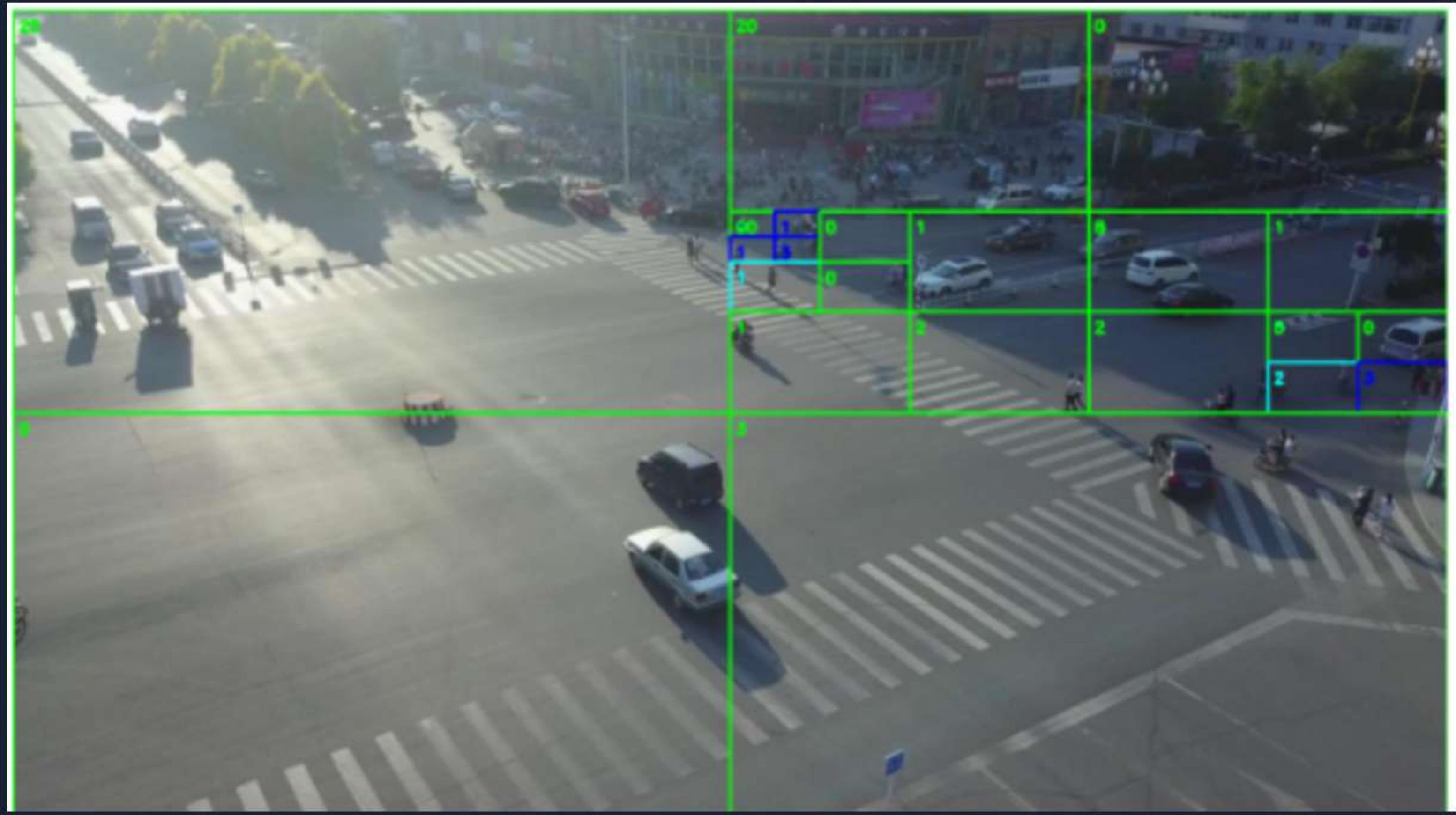
0.814
F1-Score
Balanced performance metric

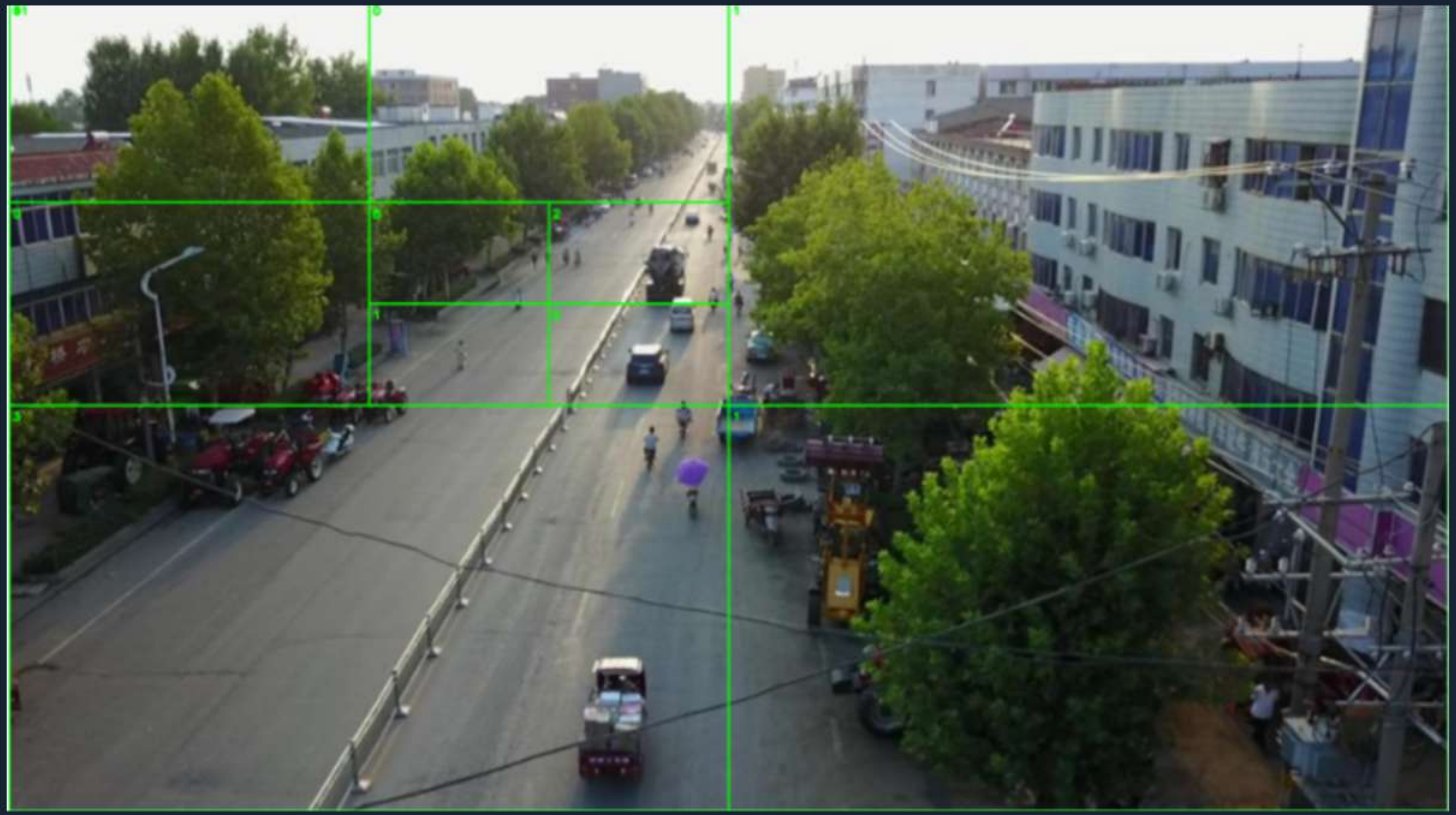
0.847
Precision
High Accuracy in person detection

0.782
Recall
Effective Identification Rate

0.791
map@0.5
Strong Detection Confidance

```
[6]  ✓ 0.0s
...
     Precision: 0.847
     F1 score:  0.814
     Recall:    0.782
     Accuracy: 0.821
     Overall Score: 0.801
     mAP@0.5:  0.791
```





Quadtree Analysis: Intelligent Resource Allocation

Our crowd density mapping employs **Quadtree analysis**—a methodical approach that intelligently allocates processing resources to the densest portions of monitored areas.

- Dynamic Partitioning: Recursive spatial subdivision based on density
- Depth Control: Maximum 6 levels for efficiency
- Threshold System: Automatic subdivision when density exceeds limits
- Real-time Processing: Continuous analysis at frame rate

what we deliver

- Live density heatmaps with color-coded intensity
- Automatic hotspot identification and localization
- Area-specific crowd counting
- Visual quadtree structure overlay
- processing sparse regions

Proven Capabilities:

- Identifies congested areas in real-time
- Provides quantitative density measurements
- Generates actionable spatial intelligence
- Maintains performance under varying loads

Predictive Intelligence: CMGraphs & GNN Analysis

The most innovative component of our framework combines real-time crowd data with static venue information to build sequential **Crowd Mobility Graphs (CMGraphs)**.

Data Integration: Fuse real-time crowd detection with venue floor plans and spatial constraints

Graph Construction: Build temporal sequence of CMGraphs capturing crowd movement patterns

Pattern Learning: Graph Neural Network analyzes historical and current flow dynamics

Congestion Forecasting: Predict exactly where and when dangerous crowding will occur

Proactive Intervention: Provide crucial lead time for resource deployment and crowd redirection

Predictive Analytics: Graph Neural Networks (GNNs)

To move beyond real-time, we implemented a predictive layer using Graph Neural Networks (GNNs) operating on Crowd Mobility Graphs (CMGraphs).

CMGraphs

A grid-based spatial representation (100px resolution) maintains a 30-frame temporal sequence. Edge weights represent crowd flow patterns.

GNN Architecture

A 3-layer GNN uses node features (position, count, density) to predict future crowd density for each grid cell up to 10 frames ahead.

Forecasting Output

The model predicts emerging congestion hotspots before they physically materialize, enabling true preventive action.

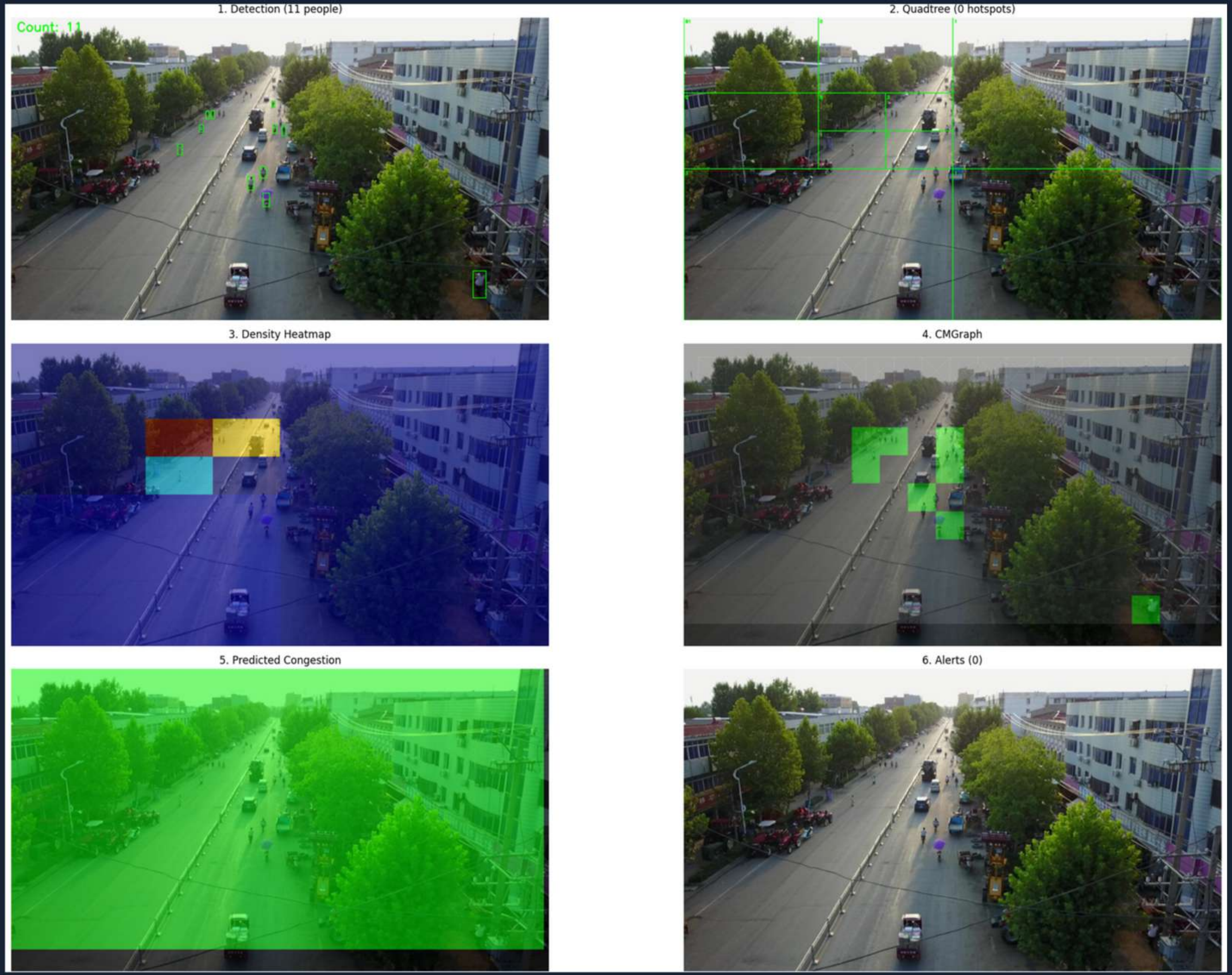
The Integrated Alert System

The final operational layer integrates both real-time density checks and GNN predictive forecasts into a robust, multi-level alert framework.

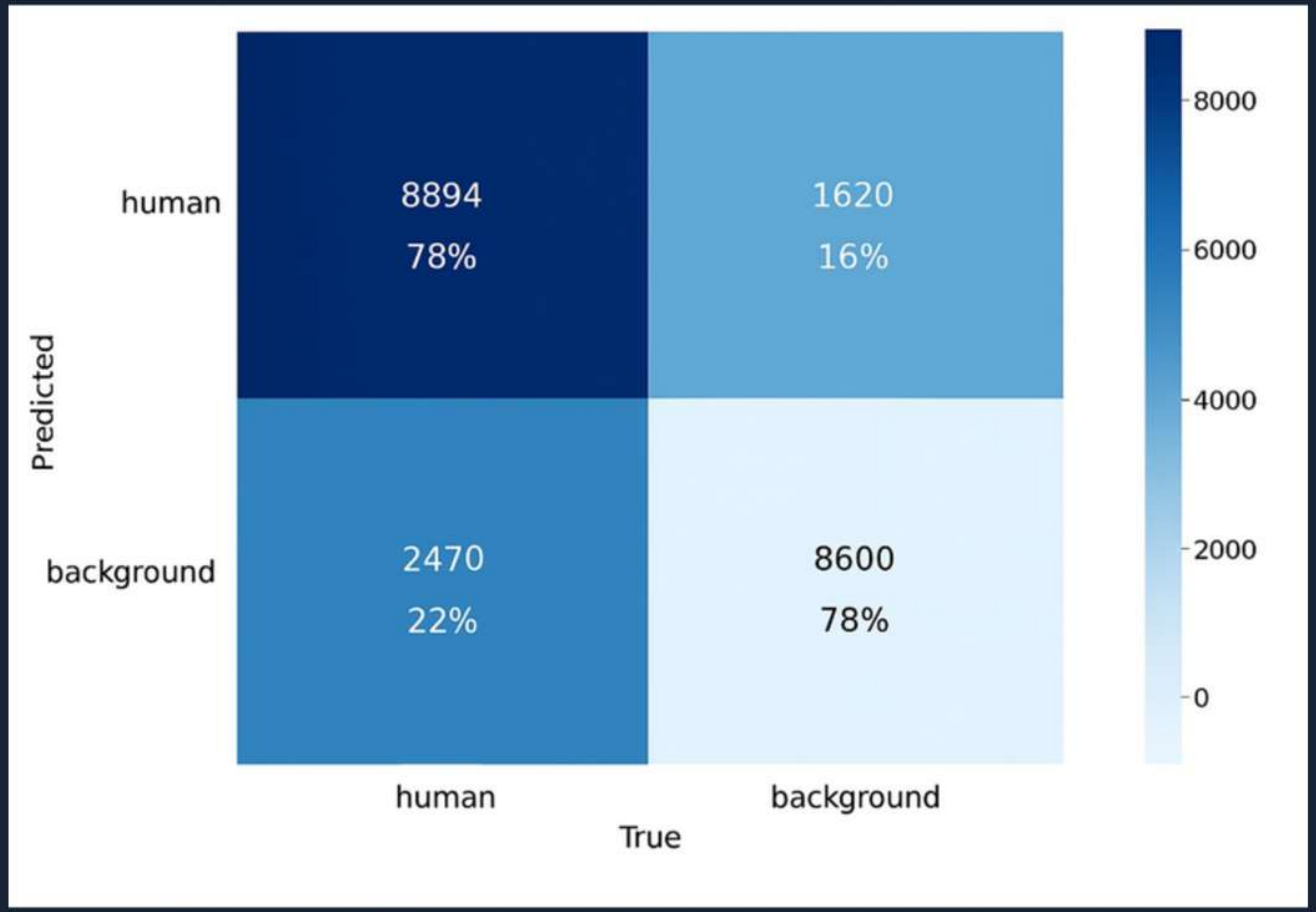
Alert Type	Description
CRITICAL_DENSITY	Real-time threshold breaches (e.g., >10 people/area), requiring immediate attention.
PREDICTED_CONGESTION	GNN-forecasted future hotspots in the next 5-10 frames, recommending preventive action.
HIGH_CROWD_COUNT	Overall venue capacity warnings (e.g., >100 people in total monitored area).

Severity Classification: HIGH alerts demand immediate intervention (e.g., rerouting) while **MEDIUM** alerts suggest preventive actions (e.g., opening secondary exits).

Results



1. Detection (11 people)
Count: 11
2. Quadtree (0 hotspots)
3. Density Heatmap
4. CMGraph
5. Predicted Congestion
6. Alerts (0)



human
background
Predicted
8894
78%
1620
16%
2470
22%
8600
78%
human
background
True
8000
6000
4000
2000
0

Conclusion

We have developed a proactive crowd management system that transforms public safety from reactive monitoring to preventive intervention. Our implemented solution delivers real-time detection, spatial intelligence, and accurate congestion forecasting with proven 82.1% prediction accuracy. This represents a significant advancement in crowd safety technology, directly contributing to safer urban environments and supporting SDG 11 objectives for sustainable cities.

Vs. Manual Monitoring

Automated, continuous, and objective analysis removes human subjectivity and fatigue.

Vs. Basic Counters

Provides deep spatial and temporal intelligence (where and when congestion will occur), not just total counts.

Technical Innovation

The first operational pipeline to fully integrate YOLO detection, Quadtree optimization, and GNN forecasting.