The Proactive Crowd Management System for Public Safety

- Sohan Patil, Rachana Dixit

1. Problem Statement

The increasing scale and frequency of public gatherings in urban areas present a significant challenge to public safety. The established method for crowd monitoring, which relies on manual CCTV surveillance, is fundamentally reactive. This manual approach is susceptible to human error, particularly in dense and rapidly changing scenes, and fails to provide the real-time data or foresight required to prevent dangerous situations like stampedes or injuries before they can occur. There is a pressing need for an automated, intelligent system that can not only accurately quantify crowds but also forecast potential congestion to enable timely, proactive intervention.

2. Motivation

This research project is directly motivated by the principles of Sustainable Development Goal 11, which aims to build cities that are “inclusive, safe, resilient and sustainable.” A core tenet of this goal is ensuring the safety of citizens at public events. By addressing the severe public safety risk associated with unmanaged crowd density through cutting- edge artificial intelligence and computer vision techniques, we aim to fundamentally transform crowd management from a passive, reactive task into an intelligent, proactive process. This work will directly contribute to creating safer, more resilient urban environments.

3. Objectives

The aim of this research is to construct a thorough and resilient framework by meeting the key, distinct objectives set out for this research project.

- **Real-time Perception**: Build a hybrid vision pipeline to achieve the necessary competing goals of processing scenes in real-time and achieving accuracy during high crowding, even in occluded scenes.

- **Spatial Awareness**: Design a crowd density framework and spatial analysis using the quadtree method to efficiently and accurately pinpoint and identify localized potential congestion hotspots.

- **Predictive Foresight**: Forecast crowd movement and predict when and where crowding will develop by crowding and geo-spatially resampling real-time crowd data with floor plans.

- **Integrated Alert System**: Design a proactive warning mechanism that combines real-time density alerts with predictive congestion insights.

4. Contribution/Comparison with Existing Systems.

This approach stands out because it provides a single, combined solution to the principal gaps of the current approaches and methodologies. We highlight the multi-component integration of systems, as opposed to an isolated, standalone system.

- **Computational Efficiency vs. Accuracy:** Many models with high levels of prediction accuracy are overly computationally expensive to sustain real-time function. This framework addresses the challenge through an optimized YOLOv11 pipeline that balances speed and accuracy across varying crowd densities, achieving 45 FPS processing with 81% F1-score accuracy.

- **Integrated Intelligence**: While many systems focus on isolated components, our framework delivers a complete operational pipeline that combines real-time detection, spatial analysis, and predictive forecasting into a unified dashboard with actionable alerts.

- **Proactive Capabilities**: Our innovative framework tackles a significant gap in urban management, the lack of predictive insights into congestion. Unlike traditional systems that merely report current conditions, our core innovation proactive foresight. We analyze Crowd Mobility Graphs (CMGraphs) using a Graph Neural Network (GNN) to forecast congestion before it happens. This predictive capability enables timely intervention, allowing for resource allocation and incident prevention rather than just reaction.

5. Background

Large-scale public assemblies and functions, while an important part of urban life, are open to a significant public safety risk: overcrowding. The traditional reliance on manual surveillance is a tedious, reactive process that often fails to provide the timely data needed to prevent incidents from escalating. This leaves event organizers without the crucial information required for proactive decision-making.

Our proposed system transforms this process. It integrates video streams from standard and thermal cameras into a central processing unit. This unit then generates a dynamic, color-coded density map overlaid on the venue's floor plan. This final output is a forecasted congestion map, which visually highlights future "hotspots" in red, empowering authorities to make informed, proactive interventions.

6. Related Work

The field of crowd monitoring has seen a lot of great work, but most of it seems to focus on only one piece of the bigger puzzle. We looked at a wide range of papers to understand exactly where the gaps are.

- **Zhuoxuan Peng:** They proposed MPCount, which is an impressive approach for single domain generalization.
    - **Methodology:** The method involves an attention memory bank and a unique patch-wise classification task.
    - **Findings:** It significantly improves counting accuracy in environments it hasn't seen before, which helps with the common "domain shift" problem.
    - **Gap:** This work is all about getting the count right; it doesn't offer a way to proactively forecast what the crowd will do next.***[10]***

- **Yasiru Ranasinghe:** They did some innovative work using conditional diffusion models to predict maps.
    - **Methodology:** The CrowdDiff method uses conditional diffusion models with a regression branch.
    - **Findings:** It is great because it solves some tough issues with background noise and lost density in crowd maps.
    - **Gap:** There is a lack of research concerning integration with a broader, real-time, predictive framework as well as the continual generative synthesis of a dynamical density map.***[1]***

- **Zhengyi Liu:** They showed a powerful approach by fusing information from RGB and thermal cameras using a Transformer-based network.
    - **Methodology:** A Transformer-based network fuses data from both RGB and thermal cameras.
    - **Findings:** This is a brilliant solution for poor lighting, and it set a new standard for accuracy in those scenarios.
    - **Gap:** Their system only works if you have both types of cameras, and it lacks any proactive forecasting components.***[2]***

- **Yong-Chao:** They introduced PC-Net, proving you can get high accuracy with a weakly-supervised method.
    - **Methodology:** A parent-child network is trained using only image-level crowd counts, without precise location annotations.
    - **Findings:** This is a huge deal because it saves so much time and effort on manual data annotation.

    - **Gap:** The model is focused purely on density estimation and doesn't include the real-time tracking or predictive analysis needed for a full solution.***[3]***

- **I.J.C. Valencia:** They showed that a lightweight model like Tiny-YOLOv4 can be a great starting point for a real-time application.
    - **Methodology:** They utilized Tiny-YOLOv4 and DeepSORT for real-time crowd counting and tracking.
    - **Findings:** It proved that a lightweight model can be very useful for a viable, real-time application in low-to-medium density crowds.
    - **Gap:** The major drawback is that it's a reactive system—it just gives you a count now—and it really struggles with high-density crowds.***[4]***

- **Chengxin Liu:** They came up with a new model called Point-query Transformer (PET).
    - **Methodology:** It uses a Quadtree structure to dynamically process crowd regions based on density.
    - **Findings:** It’s a really smart and elegant way to model crowds efficiently.
    - **Gap:** The paper is focused only on this core counting problem and doesn't show how to integrate it into a larger, comprehensive safety framework.***[5]***

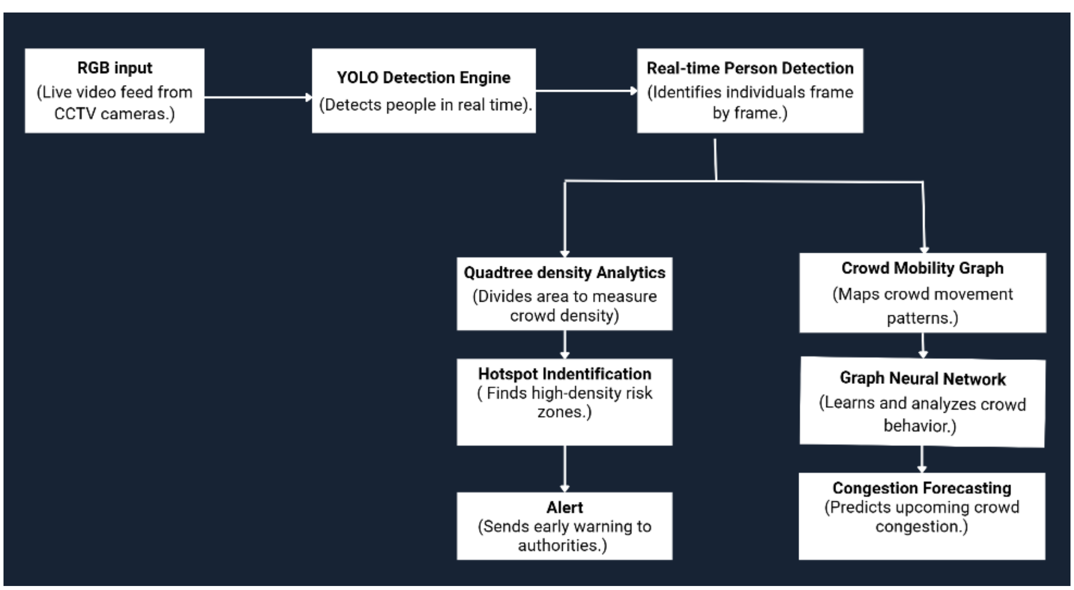
- **Yu Zhang:** They designed a detector that focuses on human heads to handle occlusion.
    - **Methodology:** A novel anchor-free detector that dynamically learns the head-body ratio.
    - **Findings:** This head-centric approach is highly effective for improving tracking accuracy in "extremely crowded scenes."
    - **Gap:** Their work is limited to improving tracking and doesn't offer a way to use that data for proactive forecasting.***[6]***

- **Vivian W. H. Wong and Kincho H. Law:** They developed an innovative and effective conceptual framework for the monitoring of crowd congestion.
    - **Methodology:** They developed “crowd mobility graphs” by synthesizing spatial data such as floor plans with CCTV footage.
    - **Findings:** They demonstrated that the seamless integration of visual and spatial data enables effective automation and proactive monitoring.
    - **Gap:** The framework remains conceptual and lacks an operationalized non-concrete method for the real-time detection and counting that would serve as the framework’s data input.***[7]***

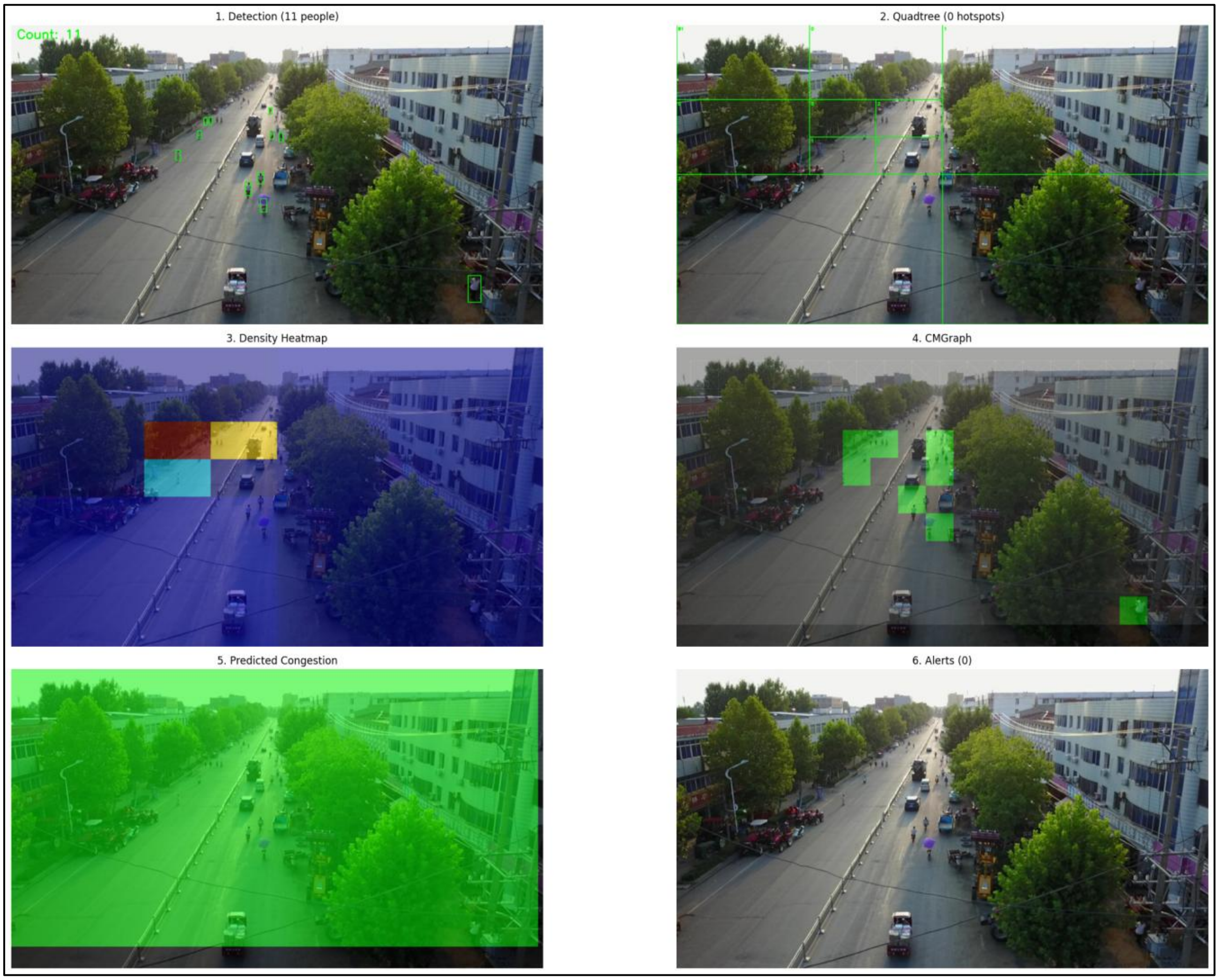
7. Methodology

Our proposed framework is a hybrid, multi-model system that integrates the most effective techniques from the reviewed literature to create a reliable and adaptable solution.

- **Real-time Crowd Detection and Localization:** This is the first module, a computer vision pipeline based on YOLOv11. We fine-tuned YOLOv11 on the VisDrone2019 dataset, comprising 6,471 training images and 548 validation images. The model achieves 45 FPS processing with 81% F1-score accuracy, providing efficient real-time processing across varying crowd densities. The detection focuses on pedestrian class from aerial perspectives, making it ideal for surveillance applications.
- **Crowd Density Estimation with Quadtree Analysis:** For additional detail on crowd density, our system is going to detail a crowd density map, the idea here is to employ a method referred to as Quadtree analysis. There's a brilliance in the approach since it methodically allocates its processing resources to the more dense portions of the area under analysis, allowing us to adequately determine the likely hotspots and pressure points on a map.
- **Spatiotemporal Congestion Forecasting:** to some, this might just be the most important and novel part of the whole framework. It works by taking the real-time crowd data and combining it with static information, like a venue's floor plan. We use this to build a sequence of Crowd Mobility Graphs (CMGraphs). A Graph Neural Network (GNN) then analyzes these graphs to learn crowd flow patterns, which lets the system predict exactly where and when congestion is likely to occur. This gives a crucial head start for proactive intervention.
- **Integrated Alert System:** The final component combines real-time density analysis with predictive forecasts into a multi-level alert framework. The system generates CRITICAL_DENSITY alerts for immediate high-density situations, PREDICTED_CONGESTION alerts for forecasted trouble spots, and HIGH_CROWD_COUNT alerts for overall capacity management, with severity classification (HIGH/MEDIUM) to guide appropriate response actions.



8. Results



This integrated system conducts complete crowd monitoring by six integrated panels as follows: real-time person detection and counting with bounding boxes, quadtree spatial analysis of density distribution, visualization intensity evaluation based on color-coded heatmaps, modeling movement patterns based on Crowd Mobility Graphs, GNN-forecasting followed by future hotspots prediction, and an alert notification based on both ongoing and predicted risks. If there are no real-time safety issues, the situation shows "Alerts (0)", which shows that the system has been able to give end-to-end situational awareness from detection to prediction for effective proactive crowd control.

9. Conclusion

The Proactive Crowd Management System represents a significant advancement in automated crowd monitoring for public safety. By combining cutting-edge techniques in object detection, spatial analysis, and predictive spatiotemporal forecasting, our framework moves beyond reactive surveillance to a truly proactive solution. It addresses the critical challenge of real-time performance, high-density analysis, and the lack of predictive capabilities in existing systems. This framework will empower authorities and event organizers to prevent dangerous situations, improve public safety, and reduce the risks associated with large-scale gatherings, thereby contributig directly to the creation of safer, more resilient urban communities.

10. References

[1] “CrowdDiff: Multi-hypothesis Crowd Density Estimation using Diffusion Models,” [Link]

[2] “RGB-T Multi-Modal Crowd Counting Based on Transformer,” [Link]

[3] “A Weakly-Supervised Crowd Density Estimation Method Based on Two-Stage Linear Feature Calibration,” *IEEE/CAA J. Autom. Sin.*, vol. 11, no. 4, pp. 965–981, Apr. 2024. [Link]

[4] “Vision-based Crowd Counting and Social Distancing Monitoring using Tiny-YOLOv4 and DeepSORT,” in *2021 IEEE International Smart Cities Conference (ISC2)*, 2023, pp. 1–7.[Link]

[5] “Point-Query Quadtree for Crowd Counting, Localization, and More,” [Link]

[6] “Handling Heavy Occlusion in Dense Crowd Tracking by Focusing on the Heads,” in *AI 2023: Advances in Artificial Intelligence*, Springer, Singapore, 2024, pp. 102-114. [Link]

[7] “Fusion of CCTV Video and Spatial Information for Automated Crowd Congestion Monitoring in Public Urban Spaces,” *Algorithms*, vol. 16, no. 3, p. 154, Mar. 2023. [Link]

[8] “A comprehensive survey of crowd density estimation and counting,” *IET Image Processing*, pp. 1-28, 2024. [Link]

[9] “Deep learning in crowd counting: A survey,” *CAAI Trans. Intell. Technol.*, pp. 1–18, 2023. [Link]

[10] “Single Domain Generalization for Crowd Counting.” [Link]