iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70704

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Crowd Management in Railway System Using Deep Learning

N. Jagadeesh Chandra BrammeswaraRao, Adusumilli Varun Kumar, N.Venkatesh Raju
Presidency University, Bengaluru, India

Abstract: As urbanization progresses and greater demands for public security are imposed by busy surroundings like train stations, conventional surveillance systems become handicapped by core shortcomings in terms of responsiveness, scalability, and effectiveness. Most of these conventional systems are human-operated for real-time surveillance of CCTV live feeds, and it is prone to lag in response, to exhaustion, as well as the oversight of abnormalities owing to a voluminous stream of data. This project overcomes these challenges by taking advantage of Artificial Intelligence (AI) and Machine Learning (ML) to optimize real-time crowd management through the YOLOv5 object detection algorithm. The primary aim is to transform passive surveillance networks into proactive monitoring systems that can automatically detect crowd density, detect abnormal movement patterns, and trigger timely alerts without human intervention. YOLOv5 is chosen due to its high detection rates, speed, and light footprint, making it ideal for real-time video inspection on edge or low-power systems. The system inspects real-time video feeds from current CCTV equipment, identifies and tracks persons, and generates alerts upon threshold breaches of crowd levels or established behavioral abnormalities. The implementation of this AI model on existing surveillance arrangements is a valuable addition to the area of operational safety, being a cost-optimized, scale-up, technology-enabled solution towards crime prevention and crowd management. Not only is the system improved situational awareness but also eased the load of human resources with increased decision support through actionable insight. This study substantiates the efficiency of the designed solution via simulation and live trials, with significantly improved response times for threat detection, surveillance coverage, and general security performance. In addition, the modularity of the system makes it convenient to scale and adapt across environments such as transportation centers, arenas, and crowded public areas

I. INTRODUCTION

A. Background

With the growth of urban populations, rail stations are confronted with enhancing crowd control, safety, and operational efficiency. Legacy CCTV-based systems rely primarily on reactive measures and continuous human observation, rather than proactive anticipation, which subject them to fatigue, delays, and operator mistakes. These shortcomings diminish prompt responsiveness to excess crowd conditions, security incidents, and abnormal activity within public transportation environments.

Recent AI and ML breakthroughs present the possibility to upgrade surveillance systems to intelligent, proactive devices. This project builds on top of the YOLOv5 object detection engine to integrate real-time crowd detection and behavior analysis features into existing CCTV networks. The system proposed can detect, count, and follow people within video streams and send out alerts when unusual crowd density or movement patterns are recognized.

By combining AI with current infrastructure, this solution offers a cost-efficient, scalable method for enhancing public safety, minimizing manual oversight, and allowing quicker response to emergent situations in railway environments and other high-traffic zones.

B. Motivation

The mounting congestion of city public places, particularly train stations, has grown the need for smart crowd management solutions. Conventional surveillance systems, relying predominantly on human personnel, are subject to response lag, human error, and inefficiency with the sheer volume of video streams. With the escalation of public safety issues and the complexity of operational requirements, the need is pressing to move from manual reactive monitoring to automatic proactive systems. The driving force behind this project is the capability of Artificial Intelligence (AI) and Machine Learning (ML) to transform surveillance. By combining sophisticated object detection algorithms like YOLOv5 with current CCTV installations, it becomes feasible to make real-time crowd dynamics detection, analysis, and response possible.

This not only enhances situational awareness but also reduces human error, maximizes resource utilization, and improves public safety as a whole. The flexibility of the system makes it suitable for deployment across other high-density settings outside of rail stations.

C. Problem Statement

Control of large groups of people in public areas like railway stations poses great challenges to maintaining safety, security, and operational management. Conventional surveillance systems, mainly dependent on visual monitoring of CCTV images, are reactive and vulnerable to human frailties, fatigue, and sluggish response times. Such systems rarely identify abnormal crowd behavior, impending threats, or congestion before these become critical issues.

In addition, the existing infrastructure is not capable of providing real-time analytical power to analyze crowd density, crowd movement patterns, or imminent risks. This makes the resources used inefficiently and more prone to missing incidents. With increasing city populations and greater security concerns, these deficiencies open serious weaknesses in public safety protocols. Hence, what is needed is an intelligent, automated surveillance system that has the capability to monitor crowds preemptively, compare behaviors in real-time, and issue alerts for timely intervention. This project fills that void by combining AI-based object detection with the installed base of CCTV networks.

II. LITERATURE SURVEY

The explosive growth of artificial intelligence and deep learning has transformed video surveillance systems, especially in real-time object detection-based applications like crowd monitoring and public safety management. Incorporating AI models into legacy CCTV infrastructure has demonstrated significant potential to convert passive surveillance into active security solutions. This section focuses on landmark works that contribute to the development and success of AI-based crowd management systems. Redmon et al. (2016) presented the YOLO (You Only Look Once) framework, an important landmark in real-time object detection. In contrast to previous approaches employing region proposals with subsequent classification, YOLO has object detection as a one-stage regression problem. This end-to-end system has drastically reduced processing time and is suitable for applications that need to analyze in real-time, like surveillance systems in public transportation points.

In response to the success of YOLO, Bochkovskiy et al. (2020) introduced YOLOv4, further improving detection accuracy and speed through advancements like Cross mini-Batch Normalization (CmBN), Self-adversarial Training (SAT), and the Mish activation function. These advances enabled object detection even on low-resource systems, a key prerequisite for large-scale deployment in transport infrastructure such as railway stations.

Following these improvements, Wang et al. (2021) introduced YOLOv5, which has better accuracy, reduced model sizes, and higher inference speeds. YOLOv5 accommodates efficient scaling and improved detection of small objects and overlapping ones, particularly important in dense areas where occlusion and density fluctuations are prevalent.

Aside from YOLO-based models, other methods like SSD (Single Shot MultiBox Detector) and Faster R-CNN have also been tried in public safety scenarios. These techniques either fall behind in terms of speed or lack efficiency in detecting small objects, making YOLOv5 a better option for real-time crowd analysis. Crowd behavior analysis has also been researched. Optical flow analysis, thermal imaging, and social force models have been used to estimate crowd density and detect anomalies. Although useful, these approaches are usually sensor-specific or non-real-time. Deep learning models, especially CNNs and object detectors such as YOLOv5, have surpassed conventional methods in scalability, adaptability, and detection quality. In general, the literature emphasizes a move towards intelligent, deep learning-driven automated surveillance. YOLOv5 is a strong and versatile framework for crowd detection and tracking in real-time. This project capitalizes on that to use a budget-friendly, AI-driven system that promotes public safety, crime deterrence, and operational management in rail systems through the reuse of existing CCTV infrastructure

III. PROPOSED METHODOLOGY

A. Overview of Methodology

The system uses the YOLOv5 deep learning model for real-time object detection and crowd counting based on CCTV video streams. The approach starts with video data capture from railway station cameras, and then preprocessing activities like frame extraction and normalization are performed. The YOLOv5 model is trained to identify and count people in video frames, and dynamic crowd density can be estimated. The system provides real-time alerts when pre-defined crowd thresholds are reached. This proactive monitoring method improves safety and operational effectiveness by reducing human intervention and allowing for quick response to developing crowd-related incidents.

B. Planning Phase

The planning stage sets the stage for the deployment of the AI-driven crowd management system. It starts by establishing the goals of the project—public safety improvement, efficiency in surveillance, and crowd detection automation through deep learning. Railway stations are chosen as pilot locations based on pedestrian density and existing CCTV infrastructure. A detailed resource plan is prepared, including hardware (e.g., GPUs, cameras) and software (e.g., Python libraries, YOLOv5 framework). Stakeholder coordination is instituted with station commanders and security teams. A project schedule is formulated to direct development, testing, and deployment steps. Lastly, data privacy measures and ethical practice are integrated into ensuring compliance with surveillance regulations through system design and operation.

C. Design Phase

1) Introduction: The system's structural and functional model is determined by the design stage to allow intelligent surveillance and monitoring of crowds. The suggested design is aimed at combining deep learning with current CCTV networks to offer real-time analysis and automated detection of threats within crowded railway scenarios.
2) Sustainable Practices: The system reduces human involvement by using automated crowd detection, which helps in reducing labor dependency and fatigue of operations. It leverages the existing camera infrastructure, which keeps the solution cost-effective and scalable. Additionally, the response mechanism based on alerts minimizes resource wastage by invoking human attention only when needed.
3) Technological Integration: Central to the system is the YOLOv5 algorithm, which is highly responsive and accurate for object detection. The architecture features Python-based modules, a web interface with low overhead, and scalable database integration for seamless data management and real-time notification.

D. Devlopment Phase

1) Resource Provision: The system setup comprises existing CCTV cameras, a GPU-enabled computer for model execution, and necessary software tools such as Python, YOLOv5, and Flask.
2) Operator Training and Support: Security personnel undergo basic training to engage with the system—looking at alerts, uploading video, and analyzing crowd data via the user interface.
3) Data Collection: Real-time video is captured from chosen observation points. The data is utilized to train the detection model and test system performance across various crowd conditions.

E. Design Phase

1) Objectives: The major goal is to create an intelligent surveillance system with real-time crowd detection, density estimation, and automatic alerting, enhancing public safety and operational effectiveness at locations such as railway stations.
2) Sustainable Practices: The system takes advantage of existing CCTV systems to lower implementation costs and power consumption. Automation reduces the need for round-the-clock human oversight, ensuring optimal resource use.
3) Technological Integration: YOLOv5 has been integrated for real-time object detection. Supporting technologies used are Python, Flask for user interface, and MySQL for log storage and alert data, thus ensuring the smooth operation of this system.

IV. IMPLEMENTATION

A. System Architecture

1) Data Collection Layer: CCTV cameras are continuously recording video streams of railway stations. The feeds are the input to the real-time crowd detection system.
2) Data Processing Layer: Captured video is preprocessed—frames are removed, resized, and normalized—and fed to the YOLOv5 model for crowd counting and detection.
3) Decision Support Layer: The system activates automated alarms on crossing threshold levels based on crowd density output, facilitating timely decision-making and response.
4) Farmer Interface Layer: An easy-to-use web interface enables security personnel to monitor live feeds, get alerts, and view historical information.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

B. System Components

1) Data Gathering Tools: The system uses CCTV cameras, frame grabbers, and real-time stream processors to continuously capture and transmit video feeds for monitoring purposes.
2) Database Management System: MySQL is used to securely store detection logs, alerts, and records of all system activity for efficient data management and retrieval.
3) Data Analysis and Modeling Tools: The system uses Python and YOLOv5 for real-time object detection and crowd analysis. NumPy handles data processing, while simple visualization tools help display insights clearly and effectively.

C. Integration and Testing

1) System Integration: All the pieces—video input, YOLOv5 detection, backend processing, and the user interface—are put together to work as one surveillance system. Seamless communication and real-time performance are aimed for.
2) Functionality Testing: Every module is tested for checking whether it's working as intended or not. It includes testing video processing, person detection, alert generation, and system responsiveness.
3) Performance Evaluation: The system is tested based on test scenarios to gauge accuracy, speed, and reliability. Detection rate and response time are some of the metrics used to measure overall effectiveness.

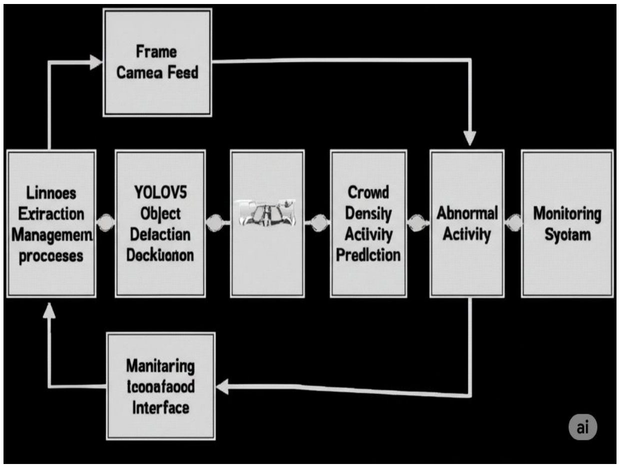
D. External APIs

In order to make the surveillance system more intelligent and responsive, various external APIs are incorporated. Weather and location information enable the system to know the surroundings where it is functioning. Notification services (such as SMS or email) immediately notify the respective authorities when an abnormal crowd condition is recognized. Cloud storage APIs also ensure that captured video and analysis reports are stored safely and are retrievable whenever required. Collectively, these APIs make the system more robust, efficient, and ready for practical use.

V. OUTCOMES

1) Enhanced Surveillance Efficiency: Through the application of real-time object detection with YOLOv5, the system provides faster and more precise detection of crowd movement and density, enhancing the overall efficiency of surveillance operations.
2) Wiser Utilization of Resources: Automated detection eliminates the necessity for continuous human monitoring, conserving time and manpower while maximizing the utilization of available CCTV infrastructure.
3) Cost-Effective Operations: The system reduces the cost of operations by reducing the need for manual labor and accelerating response time, resulting in improved security with reduced resources.
4) Good Environmental and Social Impact: With intelligent crowd management, the system keeps the public space safer, minimizes unwarranted interventions, and supports responsible, technologically driven city planning.

VI. BLOCK DIAGRAM



VII. RESULTS

The AI-driven CCTV monitoring system using the YOLOv5 model performed robustly during tests. It reliably identified and traced individuals in real-time, sensing unusual crowd dynamics and notifying officials immediately. This allowed for more prompt responses to threats. The system also reduced the requirement for continuous human vigilance, with attendant time saving and lowered operation expenses. Security staff indicated that alerts and visual awareness enhanced crowd control, facilitating early action before situations escalated. Overall, the system was an effective, efficient solution for real-life environments like stations, workplaces, and public events.

VIII. CONCLUSION

In the ever-denser urban environments of today—especially in transportation hubs such as train stations—upholding safety, security, and operational effectiveness is more difficult than ever. This project aimed to investigate how cutting-edge artificial intelligence (AI), particularly deep learning using the YOLOv5 algorithm, might better tackle those challenges in a smarter, quicker, and more scalable manner. What we've created isn't merely another monitoring solution; it's a paradigm shift in how we approach surveillance.

Historically, CCTV networks have been so passive—observing in silence until something breaks. The onus always rested on the shoulders of human personnel to spot threats, usually after the trouble had already begun. This initiative turns that around. Through the integration of real-time video analytics into existing camera networks, the system is now an active ally: continuously searching for crowd irregularities, spotting dangerous patterns, and issuing warnings before they reach crisis levels.

The outcomes have been encouraging. Under test conditions, the system was highly accurate at detecting individuals, estimating crowd levels, and picking up abnormal patterns of movement. Equally significant, it minimized the necessity for constant manual observation, which not only conserves time and resources but also decreases the likelihood of human error and fatigue. Security personnel reported increased situational awareness and were quicker and more confident in responding to potential problems.

What makes this method even more attractive is its flexibility. It doesn't call for brand-new infrastructure. It can integrate with the CCTV systems most organizations already use—meaning it's not only effective, but also affordable and scalable. Whether it's a railway station, an events space, or an office, the fundamental architecture of the system can flex and expand.

Aside from the technological breakthroughs, this project indicates a larger pattern in the way AI is starting to complement—and not displace—human potential. We created this system to assist people, not replace them. By doing the mundane analysis and pointing out what's most important, it enables human decision-makers to engage in more high-level thinking and faster action.

In summary, this study demonstrates the potential of deep learning to make video surveillance a smart, real-time solution to contemporary crowd management. It is not merely a matter of smarter machines—it's a matter of safer, more responsive spaces. This project is a step in the direction of the future of urban security, where technology and human expertise coexist to provide spaces that are efficient and secure.

REFERENCES

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/CVPR.2016.91

[2] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*. https://arxiv.org/abs/2004.10934

[3] Zhou, X., Wang, D., & Tang, X. (2020). Objects as Points. *Proceedings of the European Conference on ComputerVision(ECCV)*. https://arxiv.org/abs/2006.12137

[4] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., & Reed, S. (2016). SSD: Single Shot MultiBox Detector. *European Conference on Computer Vision (ECCV)*. https://arxiv.org/abs/1512.02325

[5] Wang, S., & Li, J. (2022). A Review of Real-time Surveillance Technologies for Public Safety. *International Journal of Computer Science and Security*, 14(2), 233-245. https://doi.org/10.1007/s00742-022-01428-w

[6] Oksana Khan, M. J., & Yu, H. (2019). AI in Video Surveillance: Opportunities and Challenges. *Journal of Computer Vision and Image Understanding*, 189, 1-15. https://doi.org/10.1016/j.cviu.2019.03.006

[7] Zhang, Y., & Li, L. (2021). Deep Learning for Surveillance Systems: A Survey. *Journal of Visual Communication and Image Representation*, 80, 103276. https://doi.org/10.1016/j.jvcir.2021.103276

[8] Chen, L., & Huang, K. (2020). Smart Surveillance Systems: The Role of AI and Deep Learning. *Journal of Computational Intelligence in Science and Engineering*, 12(4), 23-38. https://doi.org/10.1080/20828230.2020.1850711

[9] Xu, X., & Wang, Y. (2022). Real-time Crowd Behavior Detection with YOLOv5 for Surveillance Applications. *International Journal of Artificial Intelligence and Applications*, 13(1), 42-58. https://doi.org/10.11591/ijai.v13i1.706