IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, January 2026
Impact Factor: 7.67

Crowd Behaviour Anomaly Identification Using AI

Daivik Girish M[1], Deepak HB[2], Devika BS[3], Hamsashree NS[4]
BE,CSE, Kalpataru Institute of Technology, Tiptur, India [1-4]

Abstract: *Crowd safety is a major concern in public environments such as shopping malls, railway stations, college campuses, and large event gatherings. Delayed identification of abnormal crowd behaviour—such as sudden running, panic movements, physical violence, or group conflicts—can escalate into serious accidents, injuries, and property damage. Conventional CCTV systems depend entirely on human operators to monitor multiple screens simultaneously, making the process slow, labour-intensive, and highly prone to human oversight. This highlights the urgent need for an automated, intelligent, and reliable monitoring solution.*

This project introduces an AI-Driven Crowd Behaviour Anomaly Identification System that detects unusual and potentially harmful activities using advanced Pose Estimation and Deep Learning techniques. The system employs YOLOv8-Pose to accurately detect individuals and extract skeletal key points from each video frame. These motion-based features are further analyzed using a trained Temporal Autoencoder Model, which learns the patterns of normal human movement over time. Any deviation from these learned patterns results in a high reconstruction error, enabling the system to differentiate normal actions from suspicious or abnormal behaviour. Upon detecting an anomaly, the system highlights the affected region and automatically saves snapshot evidence for later review.

The proposed system is integrated into a user-friendly Streamlit dashboard, providing real-time frame visualization, alert notifications, and automatic snapshot saving for evidence. The design supports both uploaded videos and live monitoring setups. This project demonstrates an effective end-to-end pipeline for crowd behaviour analysis, emphasizing accuracy, real-time performance, and practical usability. Such an approach can significantly support smart surveillance systems in public places like malls, railway stations, stadiums, and metro hubs—ultimately enhancing safety and situational awareness through automated AI-based monitoring.

Keywords*:* Crowd Behaviour Analysis, Pose Estimation, Anomaly Detection, AI-Based Video Surveillance

I. INTRODUCTION

1.1 Problem Statement

Crowd abnormality detection is a challenging task due to the complexity of crowd behaviour and environmental conditions. Abnormal activities are context-dependent and may vary across different scenes. For instance, running is normal in a playground but abnormal in a shopping mall. Therefore, a robust detection system must adapt to varying conditions while maintaining real-time performance. The problem can be stated as: To design and implement an intelligent vision-based system capable of automatically detecting abnormal human behaviour in massive crowd scenes using pose estimation and deep-learning-based temporal anomaly detection. The system should process real-time video streams, analyze crowd motion and individual poses, and generate alerts whenever irregular patterns are detected.

1.2 Objective

The primary objectives of this project are as follows:

• To study and analyze various existing methodologies for crowd-behaviour and anomaly detection using deep learning and computer vision.

• To develop a robust and real-time framework that integrates pose estimation (using YOLOv8-Pose) and motion-based features for behaviour understanding.

DOI: 10.48175/IJARSCT-30701

• To train and deploy a temporal autoencoder model that can learn the normal motion patterns of individuals and detect deviations indicating abnormal activity.
• To implement a visualization dashboard using Streamlit for live monitoring, statistical graphs, and alert management.
• To evaluate the performance of the developed system based on metrics such as accuracy, precision, recall, and frame-per-second (FPS) processing rate.

1.3 Scope of Project

The project primarily focuses on vision-based crowd surveillance and real-time abnormality detection. The scope includes:
• Real-time or recorded video input from surveillance cameras.
• Detection of humans and estimation of their body key-points.
• Analysis of motion and temporal patterns to identify abnormal events.
• Real-time visualization and alert generation in the form of text, sound, and snapshot storage.
The system is non-invasive — it does not require any wearable sensors or biometric data, ensuring privacy. It can be deployed in various environment
The project focuses on:
• Real-time video analysis from either webcam or uploaded footage.
• Feature extraction using keypoint data and optical-flow motion vectors.
• Autoencoder-based anomaly detection using reconstruction-error thresholds.
• Interactive visualization via a web dashboard for monitoring and alerting.
The system can be scaled further for city-wide deployments as part of smart surveillance networks. However, the project does not cover individual identification, facial recognition, or predictive crowd control mechanisms. Its scope is limited to real-time detection and alert generation.

1.4 Project Context and Strategic Imperative

Rapid urbanization and the increasing frequency of large-scale public gatherings—such as railway commuters, festivals, stadium events, and metropolitan hubs—have significantly heightened the risk of crowd-related incidents including stampedes, panic movements, violence, and sudden evacuations. In such high-density environments, even a minor abnormal event can escalate rapidly, leading to loss of life, infrastructure damage, and public disorder. Traditional surveillance systems rely heavily on continuous human monitoring of multiple CCTV feeds, a process that is not only labor-intensive but also prone to fatigue-induced errors and delayed response.
In this context, the strategic imperative for modern surveillance infrastructure has shifted from reactive incident investigation to proactive, real-time risk prevention. The ability to automatically detect early signs of abnormal crowd behaviour—such as sudden running, unusual motion patterns, or aggressive interactions—is critical for ensuring public safety and operational readiness. This project is commissioned to address this challenge by transitioning from manual surveillance to an AI-driven intelligent monitoring framework. The core technical solution involves the development and deployment of a Machine Learning–based Crowd Behaviour Anomaly Detection System, capable of identifying deviations from normal movement patterns in real time and generating timely alerts to support rapid intervention.

II. METHODOLOGY

The methodology outlines a structured and systematic approach for video data acquisition, preprocessing, feature extraction, model training, and evaluation to achieve accurate and real-time detection of abnormal crowd behaviour.

A. Data Description and Sources

The system operates on video data captured from surveillance cameras in both indoor and outdoor public environments. The dataset consists of video sequences representing normal crowd activities such as walking, standing, and routine movement, along with abnormal scenarios including running, falling, sudden dispersal, and unusual motion patterns.
Each video frame is processed to detect multiple individuals and extract pose-based skeletal keypoints. The primary

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, January 2026
Impact Factor: 7.67

objective is to classify observed behaviour into two categories:
•Normal Behaviour
•Abnormal Behaviour
Key features used in the model span three critical dimensions of crowd analysis:
1.Human Pose and Spatial Features:
Skeletal keypoints (head, shoulders, elbows, knees, ankles) extracted using YOLOv8-Pose.
2. Motion and Temporal Dynamics:
Pose velocity vectors, frame-to-frame motion changes, and optical-flow magnitude.
3.Interaction and Crowd-Level Activity:
Collective motion intensity, sudden acceleration patterns, and temporal consistency of movements.

B. Data Preprocessing and Cleaning

To ensure reliable model performance, rigorous preprocessing was applied to the raw video data.
1. Frame Preprocessing
• Video frames were resized and normalized for consistent model input.
• Frames were converted to grayscale for optical-flow computation to reduce computational overhead.
2. Pose Normalization and Noise Reduction
• Extracted skeletal keypoints were normalized relative to body size to ensure scale invariance.
• Temporal smoothing and filtering were applied to reduce jitter caused by camera motion or detection noise.
3. Feature Transformation
• Temporal Sequencing: Pose velocities were aggregated into fixed-length temporal windows.
• Scaling: Motion features were normalized to prevent dominance of high-magnitude movements.
• Feature Fusion: Pose-based features were combined with optical-flow motion metrics for robust anomaly detection.

III. RELATED WORK AND THEORETICAL FOUNDATION

A. Rationale for Automated Crowd Anomaly Detection

Prior research establishes that delayed detection of abnormal crowd behaviour is a major contributor to large-scale disasters. Manual monitoring systems are inherently limited in dense environments where individual tracking becomes difficult. Studies emphasize the importance of proactive surveillance systems capable of identifying early deviations from normal crowd dynamics to prevent escalation.

B. Application of Machine Learning in Crowd Behaviour Analysis

Crowd behaviour modeling has evolved from handcrafted motion descriptors to advanced deep learning frameworks:
• Traditional Methods: Optical-flow histograms and trajectory clustering provided limited scalability.
• Deep Learning Models: CNNs and Autoencoders demonstrated improved performance by learning normal motion patterns directly from data.
• Pose-Based Models: Recent studies highlight that skeletal motion features offer higher interpretability and robustness in crowded scenes.
The current project adopts YOLOv8-Pose combined with a Temporal Autoencoder, leveraging both spatial and temporal learning capabilities.

C. Precision-Oriented Detection and Safety Assurance

A key challenge in real-world surveillance systems is minimizing false alarms. Excessive false positives reduce operator trust and overload response systems. Therefore, this work emphasizes precision-oriented anomaly detection, ensuring that alerts are generated only when significant deviations from learned normal patterns occur. This design choice aligns with practical safety requirements where reliability and actionability are critical.
Fig.1 illustrates the temporal variation of motion magnitude and autoencoder (AE) reconstruction error across video frames. The plotted curves represent normal crowd movement patterns, while the dotted horizontal line indicates the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, January 2026
Impact Factor: 7.67

predefined anomaly threshold. The figure demonstrates that, under normal conditions, both motion intensity and AE error remain consistently below their respective thresholds. Sudden spikes beyond these limits indicate abnormal crowd behaviour. The contrasting line colors are clearly distinguishable on screen and remain interpretable when printed in black-and-white format, ensuring effective visual analysis.

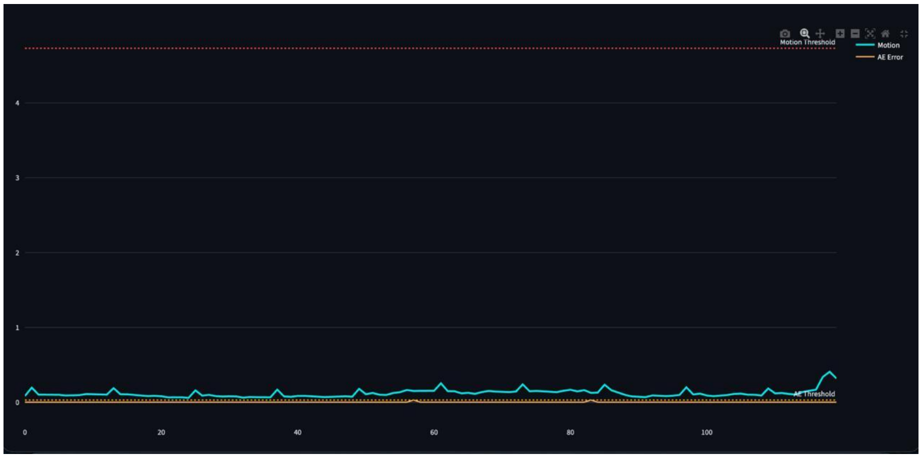

Fig. 1 Motion and AE Error Analysis – Test Scenario 1

Fig. 2 shows a sample frame from the proposed crowd behaviour anomaly detection system. Individuals detected in the scene are enclosed within bounding boxes, where green boxes indicate normal behaviour and the red box highlights an anomalous subject identified by elevated autoencoder reconstruction error. The associated AE error values are displayed above each bounding box for interpretability. The image resolution is sufficient to clearly reveal human posture, bounding box boundaries, and text annotations, ensuring all visual details remain legible both on screen and in printed hardcopy.

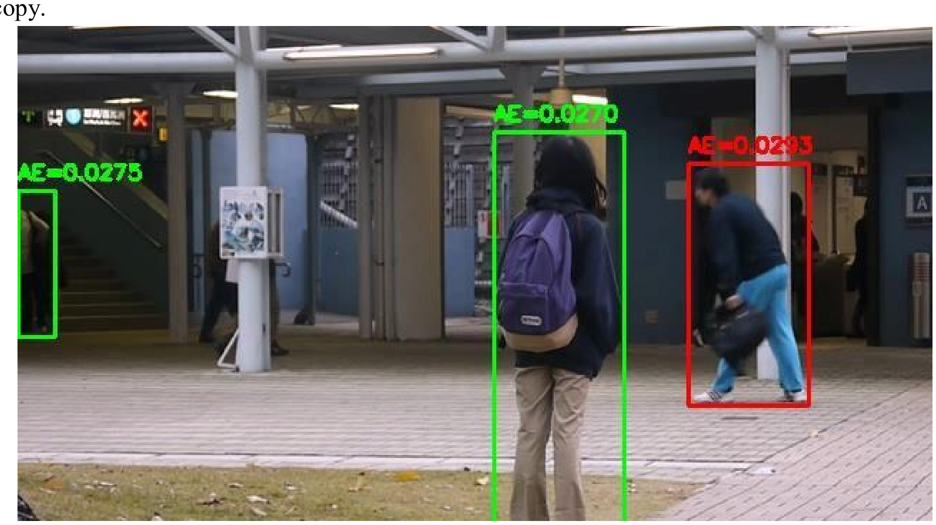

Fig. 2. Real-Time Crowd Anomaly Detection and Localization

IV. CONCLUSION

The project "Abnormal Behaviour Detection in Massive Crowd" successfully demonstrates an automated AI-driven surveillance system capable of identifying irregular and suspicious human activities in densely populated environments. By integrating YOLOv8-Pose for accurate multi-person detection and pose estimation with a Temporal Autoencoder for anomaly recognition, the system achieves real-time detection of unusual events such as running, fighting, and falling. The combined use of spatial pose information and temporal motion analysis enables effective understanding of

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, January 2026
Impact Factor: 7.67

behavioural changes over time, while optical-flow-based global motion estimation enhances reliability by detecting sudden crowd-level disturbances even when pose data is partially unclear. The proposed solution includes an intuitive Streamlit-based dashboard that provides live video visualization, anomaly statistics, motion graphs, alert logs, snapshot storage, and voice alerts, supporting rapid operator response in critical situations. Overall, the system minimizes human supervision, improves detection accuracy, and enables faster incident response, demonstrating that deep learning–based surveillance solutions significantly outperform traditional manual monitoring methods in terms of speed, scalability, and operational effectiveness for deployment in public spaces such as railway stations, airports, shopping malls, and stadiums.

V. ACKNOWLEDGMENT

We express our sincere gratitude to the Visvesvaraya Technological University (VTU) for providing us the opportunity to carry out this project as part of the academic curriculum. We are deeply thankful to the Department of Computer Science and Engineering, Kalpataru Institute of Technology (KIT), Tiptur, for providing a supportive academic environment and necessary resources throughout the course of this work.

We would like to convey our heartfelt thanks to our project guide, whose valuable guidance, constant encouragement, and constructive feedback played a crucial role in the successful completion of this project. We are also grateful to the Head of the Department, faculty members, and laboratory staff for their continuous support and cooperation.

We extend our appreciation to our friends and classmates for their collaboration and motivation during various stages of the project. Finally, we are immensely thankful to our parents and family members for their unwavering support, encouragement, and understanding, which have been a constant source of inspiration throughout this endeavor.

REFERENCES

[1]. Mehran R., Oyama A., Shah M., "Abnormal Crowd Behavior Detection using Social Force Model," CVPR 2009.

[2]. Cong Y., Yuan J., Liu J., "Sparse Reconstruction Cost for Abnormal Event Detection," CVPR 2011.

[3]. Hasan M., Choi J., Neumann J., "Learning Temporal Regularity in Video Sequences," CVPR 2016.

[4]. Ravanbakhsh M., et al., "Abnormal Event Detection in Videos using 3D Conv Autoencoder," ICIP 2017.

[5]. Ionescu R., et al., "Object-centric Auto-Encoders for Video Anomaly Detection," CVPR 2019.

[6]. Ultralytics YOLOv8 Documentation, 2024.

[7]. OpenCV Library Manual, Version 4.10.

[8]. PyTorch Documentation, Meta AI 2025.

[9]. W. Luo, W. Liu, and S. Gao, "A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework," in IEEE International Conference on Computer Vision (ICCV), 2017.

[10]. D. Xu, R. Song, and X. Wu, "An Abnormal Crowd Behavior Detection Based on Motion Feature Extraction," in Pattern Recognition Letters, vol. 44, pp. 88–95, 2014.

[11]. F. I. Bashir, A. A. Khokhar, and D. Schonfeld, "Real-Time Motion Trajectory-Based Indexing and Retrieval of Video Objects," in IEEE Transactions on Multimedia, 2005.