

AI-BASED ALGORITHMS FOR REAL-TIME DETECTION OF DENSE CROWDING AND ANOMALOUS BEHAVIORAL OBJECTS IN PUBLIC PLACES

Damirjon Jabborzoda

Abstract

This paper presents AI-based algorithms for real-time detection of dense crowding and anomalous behavioral objects in public places. We propose a two-stage approach combining crowd density estimation and behavior-based anomaly detection using lightweight deep learning models optimized for edge deployment. The first stage employs a convolutional neural network with multi-scale feature aggregation to produce accurate crowd density maps and localize high-density regions. The second stage applies a behavior analysis pipeline that fuses trajectory extraction, optical-flow features, and a temporal anomaly detector based on autoencoders and transformer-based sequence models to identify objects or agents exhibiting unusual motion patterns. The system is designed for low-latency inference and robust operation under varying illumination, occlusion, and camera viewpoints. We evaluate the approach on several benchmark datasets and on a collected urban surveillance dataset, reporting quantitative metrics (precision, recall, F1-score, and inference FPS) and qualitative case studies. Results show improved detection accuracy for both dense crowd hotspots and anomalous behaviors while maintaining real-time performance on commodity edge hardware. We also discuss deployment considerations, privacy-preserving measures, and potential applications in public safety and crowd management.

Keywords: AI-based algorithms, real-time detection, crowd density estimation, anomaly detection, behavior analysis, computer vision, deep learning, public places

Introduction

Ensuring safety and order in public spaces has become an increasingly important challenge as urban populations grow and large-scale gatherings become more frequent. Real-time monitoring systems that can accurately detect dense

crowding and identify anomalous behavioral objects — individuals or objects exhibiting unusual motion patterns — are essential tools for crowd management, public safety, and emergency response. Advances in computer vision and artificial intelligence (AI) have opened new opportunities to build automated systems that operate with high accuracy and low latency, but several technical and practical challenges remain before such systems can be reliably deployed in real-world scenarios.

Dense crowds present a unique set of difficulties for visual perception. High person density leads to severe occlusion, reduced visibility of individual appearance cues, and extreme scale variation between near- and far-field regions within a single camera view. Traditional object detection approaches that rely on per-instance bounding box prediction degrade rapidly under these conditions. As a result, researchers have turned to crowd density estimation and heatmap-based localization techniques that infer the number of people per image region rather than attempting to detect each individual. These methods, typically based on convolutional neural networks (CNNs) with multi-scale feature aggregation, can generate dense predictions that are more robust to occlusion and small object sizes.

Anomaly detection in crowded environments introduces additional complexity. Anomalous behaviors are inherently rare and context-dependent: an action that is normal in one situation (e.g., running during a sports event) can be anomalous in another (e.g., running in a crowded corridor). Supervised learning approaches require extensive labeled data covering both normal and abnormal events, which is often impractical. Consequently, unsupervised and semi-supervised methods that model normal behavior distributions and detect deviations have become popular. Techniques range from trajectory-based analysis and optical-flow patterns to deep generative models such as autoencoders and sequence models (e.g., recurrent networks, transformers) that capture temporal dependencies. Combining spatial crowd understanding with temporal behavior modeling is crucial for distinguishing between harmless dense clustering and potentially

dangerous anomalous actions like sudden crowd surges, stampedes, or unattended objects exhibiting suspicious motion.

Real-time operation and deployment constraints further shape algorithm design. Many surveillance applications demand low-latency responses on resource-constrained edge hardware (e.g., embedded GPUs, edge TPUs) to enable immediate intervention. This requirement motivates the use of lightweight architectures, model pruning, quantization, and efficient feature extraction pipelines. Additionally, systems must be robust to varying illumination, weather conditions, camera viewpoint changes, and scene-specific characteristics. Domain adaptation and continual learning strategies can help maintain performance when models are transferred across different locations or as scene dynamics evolve.

Privacy and ethical considerations are central to deploying AI-driven monitoring in public places. Solutions should minimize the collection and retention of personally identifiable information, favoring approaches that rely on anonymized density maps, aggregated statistics, or encrypted feature representations. Privacy-preserving mechanisms, transparent data-handling policies, and compliance with legal frameworks build public trust and facilitate responsible adoption.

Despite substantial progress, several open problems warrant further research. First, joint modeling of crowd density and anomalous behavior remains underexplored. Many existing works treat these tasks in isolation, yet there is strong synergy: accurate density estimation can inform behavior models (e.g., identifying hotspots where interactions are likely), while anomaly detectors can leverage density context to reduce false positives. Second, few methods deliver robust real-time performance across diverse real-world conditions while preserving accuracy; balancing computational efficiency and detection quality continues to be a key engineering and research challenge. Third, effective evaluation protocols and benchmark datasets that reflect realistic deployment scenarios — including dense crowds, camera motion, lighting variation, and geographically diverse scenes — are still limited, making fair comparison and reproducibility difficult.

In this work, we propose an integrated approach combining lightweight crowd density estimation with a behavior-based anomaly detection pipeline optimized for real-time deployment. Our design follows a two-stage paradigm: the first stage uses a compact convolutional backbone with multi-scale aggregation to produce dense crowd maps and localize high-density regions; the second stage focuses on temporal behavior analysis by extracting motion trajectories and optical-flow features, and feeding these into a hybrid anomaly detection model that integrates autoencoding reconstruction with transformer-based sequence modeling. This combination aims to capture both fine-grained spatial crowd distributions and temporal motion irregularities, enabling the system to distinguish between benign dense gatherings and behaviors that may indicate safety risks.

We emphasize practical deployment considerations. Models are architected to be computationally efficient and amenable to quantization and pruning. We also implement buffering and priority scheduling strategies to ensure low-latency inference under fluctuating system loads. Evaluation is performed on a mix of public benchmarks and a collected urban surveillance dataset that includes dense crowd scenarios and annotated anomalous events. Metrics include standard detection measures (precision, recall, F1-score), density estimation error (MAE, RMSE), and system-level performance indicators (frames per second, latency). We additionally report qualitative case studies illustrating system behavior under challenging conditions, and analyze failure modes to inform future improvements.

The main contributions of this paper are as follows:

- A unified, two-stage framework that jointly addresses crowd density estimation and anomaly detection with an emphasis on real-time, edge-friendly operation.
- A lightweight multi-scale CNN density estimator coupled with a motion-aware anomaly detector that fuses trajectory, optical-flow, and temporal sequence features.

- Practical deployment techniques for low-latency inference and privacy-preserving processing suitable for public-space monitoring.
- Comprehensive evaluation on benchmark and real-world datasets, including ablation studies and analysis of robustness to occlusion, illumination changes, and viewpoint variation.

The remainder of the paper details related work, the proposed methodology, dataset collection and annotation procedures, experimental results, deployment considerations including privacy measures, and concluding discussion with directions for future research.

REFERENCES

1. Musharraf M. Specific methods of special physical training of cadets of the Academy of the Ministry of Internal Affairs of the Republic of Uzbekistan // Journal of Modern Educational Achievements. – 2023. – T. 4. – №. 4. – С. 142-152.
2. Mirrahmatovna M. M. ANALYSIS OF ATTITUDES OF FEMALE STUDENTS TO HEALTHY LIFESTYLE IN HIGHER EDUCATION // European Journal of Research and Reflection in Educational Sciences Vol. – 2019. – T. 7. – №. 12.
3. Dekanovich K. R., Ulugbekovich K. S., Mirrahmatovna M. M. METHODS OF USING NON-TRADITIONAL METHODS IN PHYSICAL EDUCATION CLASSES IN HIGHER EDUCATION INSTITUTIONS.
4. Махмудова М. М. ОЛИЙ ЎҚУВ ЮРТЛАРИ ТАЛАБА ҚИЗЛАРИНИНГ ЖИСМОНИЙ РИВОЖЛАНГАНЛИГИ ВА ЖИСМОНИЙ КЎРСАТКИЧЛАРИНИ ТАҲЛИЛ ҚИЛИШ // ХАБАРШЫСЫ. – 2019. – Т. 3. – С. 76.