

A survey on Abnormal Detection in the Crowd

¹Padmavathi H G,² Dr. Robin Rohit Vincent

¹Department of CSE

Research Scholar

Presidency University

Mail-Id: padma.bce@gmail.com

²Associate Professor and HoD

Department of CSE

Presidency University

Bengaluru-robinrohit.vincent@presidencyuniversity.in

Abstract

This article discusses employing advanced technologies to track the actions of unusual people in busy places, e.g. weird hand and body motions, and various facial expressions, all of which represent a potential security threat spot or some danger coming ahead. The main goal of the study is to integrate smartest precautions in the areas like public places, healthcare centers, and the traffic infrastructure via the use of improved algorithms, machine learning systems, deep neural networks and computer vision. In this case, the study relies on the use of models running convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which can analyze the video data and detect the anomalies in real-time, leading to PTZ, mobilization to monitor detected abnormal activities.

The article shows that the current techniques have been explored, discussing spatial feature extraction better in CNNs and LSTM networks better in temporal sequence analysis. It also, suggests the interaction of Generative Adversarial Networks (GANs) with the optical flow technique for more accurate anomaly detection in busy places with special stress on big population areas such as the Hajj pilgrimage.

Besides, the research investigates the option of Unmanned Aerial Vehicles (UAVs) for monitoring the crowd, laying emphasis on the necessity of policy as well as ethical considerations in drone surveillance. One of the key points raised in the study is the support to the real-time anomaly detection systems by hybrid models that combine several deep learning technologies to improve the accuracy and efficiency of such systems. The evidence provided supports the argument that the improvement of these technologies is going to bring about a noticeable change in the area of public safety and security through offering flexible and concrete monitoring solutions.

INTRODUCTION

Intelligent computing plays a vital role in the automatic detection of abnormal behaviors in videos utilizing deep learning techniques. This includes using state-of-the-art algorithms, machine learning techniques, deep neural networks, and computer vision to closely evaluate and spot unusual actions in video. This technology can highly improve safety procedures in different sectors like public spaces, hospitals, and transportation. This report deploys deep learning algorithms, which are a part of machine learning, to recognize the video data patterns dealing with abnormality and normality classification. This way the system can mobilize the PTZ (Pan-Tilt-Zoom) camera to the particular location that has detected

a moving object within a scene of interest in real time. For instance, we can use deep learning models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) with computer vision technology to find anomalies in the video data produced by a closed-circuit television system. CNNs can be taught to learn spatial hierarchies of features from images, and hence they are good for tasks such as object recognition.

Anomalous behavior identification can be very important in many different contexts for several landslides. This mission is indispensable to the inclusion of safety and security in our daily lives. An instance can be given about the significance of tracking abnormal movement in crowded areas like business centers and airports causing the detection

of the charted targets to be highlighted due to potential arising issues on the runway. Furthermore, identifying standard deviations can be quite advantageous in a variety of sectors. This includes the ability to pinpoint students who may require additional support, as well as recognizing issues among patients within the spheres of healthcare and education. Hence, it is crucial to not only recognize but also participate in discussions concerning unusual behaviors, with the overall objective of fostering a community that not only values safety but is also more sensitive, adaptable, and compassionate towards the requirements of its citizens.

RELATED WORK

This study suggested applying deep learning methods. To identify abnormal behaviours and acts, such as hand gestures, body movements, and facial expressions in densely populated locations, this article has proposed using deep learning algorithms. These acts and behaviours are a sign regarding potential security and safety problems. With the use of CNN known as convolutional neural networks along with grid-based system can not only result in the increase of accuracy but also efficiency of real-time surveillance systems. By looking into the CCTV footage, we can see that the techniques can investigate the crowd behaviour, identify anomalies and also provide alerts with respect to time to avoid incidents to occur. This approach uses advanced deep learning models to get peaked accuracy while detecting different kinds of abnormal behaviour making sure of safety and security publicly. This research paper proposed the use of technique CNN known as convolutional neural networks with limitations such as high computational cost and research gap to optimize for real-time processing [1].

The paper's main argument suggests that the YOLO-based process is most effective among areas of high population density, the format of a person's physical description and the movement of an object through a system created for crowd recognition. However a method still has to deal with the problems decreasing a level of light and a poor image quality also have the possibility of leading to a situation wherein the accuracy of recognition is

only half of the correct detection. These constrictions bring home the necessity of meticulous scrutiny of ways to ensure the last ability of the model convinced the issue of both fit of the system to different light conditions and image quality. Seeing these gaps in research is of great importance in the development of accurate and reliable real-time crowd detection systems that are adept at providing high-quality performance in different environmental backgrounds. The study at hand suggests the utilization of YOLO-based Crowd Detection technique, but also outlines the several it faces due to low visibility in night light and bad image resolutions. The detective novel found that the biggest unresearched mystery was how to increase detection precision such as when there is changing light and the camera is of a different quality [2].

This investigation examines the use of Closed-Circuit Television (CCTV) footage for anomaly detection in crowded areas with specific reference to Hajj pilgrimage. In this academic paper, the scholars provide a whole system that combines optic flow techniques and Generative Adversarial Networks (GANs), which is a kind of artificial intelligence architecture aimed at identifying unusual conduct in crowds. Furthermore, authors put forward an inventive model predicting crowd density, which is related to global measures for security in Makkah. This scholarly study highlights the effectiveness of applying deep learning approaches towards spotting abnormalities in human behavior within crowded environments.

The paper ends with a call for more sophisticated research investigations meant to improve not only the accuracy but also the flexibility of these impressive anomaly detection systems. In addition, this investigation recommends adoption of optical flow approach and integration of generative adversarial networks (GANs) taking into account limitations such as differences in scale and lighting conditions, speed of object tracking, existence of background clutter, variety of objects and inconsistency in camera positions. [3].

This study provides a thorough review of the purpose of UAVs in not only monitoring but also analysing densely populated areas. The paper demonstrates the architecture of the drone,

sensors onboard, power management, and analysis algorithms important for the monitoring of the crowd. The advantages of UAVs over usual methods, like increased coverage area and flexible outlook are highlighted here. The paper not only emphasizes the need for policy debate and legal frameworks to regulate the use of drones in crowds but also the ethical and privacy issues concerning drone technology. As a result, the study emphasizes the promising interest in the utility of UAVs in public safety as well as security. This research paper proposes Unmanned Aerial Vehicles abbreviated as UAVs with limitations like privacy concerns and ethical implications of drone surveillance. The research gap is the further studies should explore the integration of drones with other emerging technologies, such as AI, blockchain, and IoT, to enhance crowd monitoring capabilities. [4].

The paper discusses the significance of intelligent video surveillance in enhancing public security through anomaly detection and classification using the Faster RCNN model with deep reinforcement learning. It emphasizes the difficulties in dynamic scenes to identify abnormal activities and the progress from handcrafted features to deep learning techniques to improve accuracy. A novel IVADC-FDRL model that unites anomaly detection with classification to achieve peak accuracy rates as compared to the traditional methods is presented in this study. The model operates surveillance videos frame-by-frame by using quicker RCNN and DQL for anomaly detection as well as classification. The results on the UCSD dataset output the dominating performance of the IVADC-FDRL model while showcasing its capability for applications such as fall detection and walkway monitoring. This research paper proposes a Faster RCNN technique along with the Deep Learning Reinforcement Model (IVAC-FDRL) with limitations like the need for extensive computational resources because of the complexity of deep learning models. The research gap is the necessity to design effective models in order to achieve peak accuracy in real-world surveillance systems. [5].

The articles review the purpose of Convolutional Neural Networks in the detection of abnormal human behaviours in surveillance. Different

machine learning methodologies like clustering, classifiers, and CNN-driven models are engaged to extract characteristics like spatial and motion to identify behaviours like falling, violence, and loitering. Techniques like Curvature Scale Space (CSS) and Fisher Vector (FV), known as integration of depth sensors, background subtraction as well as feature extraction are crucial for identifying instances of falls. The self-operating system defined in the research focuses on increasing security and safety by examining surveillance videos to detect abnormal behaviours and ease undesired events. In this study, we propose a method for CNNs using Long Short-Term Memory (LSTM) with the drawback of being unable to handle overlapping objects and requiring high computational resources due to processing multiple data streams. The research gap in respect of this paper is how to recognize abnormal behaviors in crowded scenarios involving many individuals what the existing models find it tough to deal with in an efficient manner. [6]. This document stresses crowd analysis with regard to public safety and mitigating the adverse impacts. It puts forward a system of deep learning including CNN and k-means for detecting abnormal behaviour. Video surveillance is used to evaluate parameters for detecting abnormal behaviour. The system has an accuracy rate of 80.79% in identifying abnormal acts amongst the individuals involved. This paper reviews many works on crowd behaviour analysis including those that are based on ViBe and CNN algorithms respectively. It highlights the importance of real-time systems which are not only aimed at detecting unusual activities among the crowd but also their possibility to be applied within deep learning framework. In this study, a deep learning architecture with constraints such as insufficiency in gathering reliable information on unusual behaviour in overcrowded places is presented. The research gap involves further exploration into merging multiple approaches such as incorporation of CNN and K-Means algorithms to enhance accuracy and efficiency in detection of abnormality in crowds. [7].

This document compares the suggested method to others such as SMOT, IOU, GOG, CEM, SLSTM, SORT, RLSTM and RMOT in terms of evaluation

measures like FP (False Positives), FN (False Negatives), IDS (Identity Switches) and FM (Fragmentation). Compared to the existing approaches, this proposal indicates higher levels of efficiency, accuracy as well as precision. According to statistical analysis, the proposed technique minimizes false rates while surpassing its counterparts on different performance metrics. There are detailed tables and figures backing these claims. This research paper proposed the technique of RCNN-based deep learning framework with limitations like the presence of false positives and false negatives in the detection results. The research gap of it includes improving the efficiency and accuracy of object detection methods compared to existing techniques. [8]

The discourse probes into the progress made in deep learning methodologies for intuiting abnormal human conduct in surveillance footage. It discusses the obstacles presented by a scarcity of labelled data, underscoring the necessity of vanquishing any data insufficiency to create efficient detection systems. The auditing draws attention to the noteworthy influence of deep learning architectures, including convolutional neural networks (CNN), long-short-term memory (LSTM), and gated recurrent unit (GRU), in scrutinizing abnormal behaviours spatially and temporally. Diverse research endeavours are demonstrated, displaying various strategies such as weakly supervised techniques, attention mechanisms, and pseudo-label generators to boost anomaly detection efficacy. The document also references the utilization of spatial-temporal graph convolutional networks and methodologies for object detection and tracking in identifying falls across different scenarios. The examination entailed a thorough scrutiny of pertinent literature, concentrating on recent works post-2020, and classifying them according to the supervision level in anomaly identification. In essence, the manuscript embellishes valuable point of views on prevailing tendencies, hurdles, and inventive resolutions within the domain of spotting abnormal human behaviour through deep learning in surveillance videos. The PDF's author utilized a literature review approach to comprehensively survey deep learning methods for detecting

anomalous behaviour in surveillance videos. This entailed examining prominent journals and online platforms to identify recent articles on deep learning techniques. To ensure thoroughness, respected academic websites like Web of Science, IEEE Xplore, Google Scholar, Science Direct, Scopus, ACM, and MDPI were referenced. Various keywords such as deep learning, unsupervised learning, weakly-supervised learning, semi-supervised learning, anomaly detection, and video surveillance were employed for article retrieval. Moreover, a screening process was implemented by the author to classify chosen papers into unsupervised, partially supervised, and fully supervised detection methods. The methodology also encompassed evaluating the strengths and limitations of diverse abnormal behaviour detection approaches, comparing their performances on well-known benchmark datasets, and pinpointing unresolved research areas in the discipline. [9]

The article presents a hybrid approach that uses video categorization to identify unusual behaviors of elevator users. The arrangement structured here combines a gradient flow feature fusion calibre with a sophisticated multi-scale attention mechanism so as to take over the constraints in existing frameworks. The system strived to bring about a top-1 precision of 95% on a dataset of anomalous lift behaviors, showing significant gains over the PPTSM framework. Comparative evaluations using publicly accessible datasets like UCF24, UCF101 and HMDB51 features on how much more effective the suggested framework is, with the notable accuracy gains ranging from 3.96% to 21.2%. The integrated modules' effectiveness tamps down in improving framework performance and lowering parameter complexity is validated by ablation research. The study acts as a cornerstone to the performed analysis and showcases its potential to improve elevator safety by calling attention to its dependability and flexibility. The methodologies that are encountered here encompasses deep learning with a multi-stream, deep, separable convolutional neural network for the enhancement of action recognition. A multi-scale dilated attention mechanism captures subtle abnormal behavior properties, while a gradient flow feature fusion

module is interwoven for high-level semantic and low-level detail features. Experimental evaluation on an elevator dataset with four abnormal behaviors and comparisons with advanced models on public datasets demonstrate robust generalization. Ablation studies assess the effectiveness of components like the Multi-Scale Dilated Attention and Gradient Flow Information Fusion Block in enhancing performance and reducing parameters. [10]

VARIOUS CROWD ANALYSIS TECHNIQUES

For anomaly detection in crowd analysis, Deep learning techniques have shown a massive attention because they automatically learn complex features from large datasets. Here are some of the prominent deep learning methods used for detecting anomalies in crowds.

1. Convolutional Neural Networks (CNNs)

To detect anomalies in crowd scenes of images and videos CNNs are more effective.

Autoencoders: These are neural networks used for unsupervised learning of efficient codings. A CNN-based autoencoder is trained to reconstruct normal crowd patterns. Anomalies are detected by evaluating the reconstruction error; abnormal events indicated by significant deviations.

CNN-LSTM Networks: Combining CNNs for spatial feature extraction with Long Short-Term Memory (LSTM) networks for temporal feature extraction, these networks can detect anomalies by learning both spatial and temporal patterns in crowd behaviors.

2. Recurrent Neural Networks (RNNs)

RNNs, especially LSTM and Gated Recurrent Unit (GRU) networks, are used for sequence data, making them suitable for analyzing temporal sequences in video data.

LSTM Networks: LSTMs can capture long-term dependencies and are used to model and predict normal sequences of crowd movements. Anomalies are identified when the predicted behavior significantly deviates from the observed behavior.

GRU Networks: Similar to LSTMs, GRUs are used for handling sequence data with fewer parameters, making them computationally efficient for anomaly detection in crowd sequences.

3. Generative Adversarial Networks (GANs)

GANs consist of a generator and a discriminator network competing against each other. They can be used to detect anomalies by learning the distribution of normal crowd behaviors.

AnoGAN: Anomaly detection using GANs (AnoGAN) involves training a GAN on normal crowd behavior data. During inference, anomalies are detected based on how well the generator can replicate the observed data. High reconstruction error or low discriminator confidence indicates an anomaly.

Conditional GANs (cGANs): These GANs generate data conditioned on additional information, such as the context of the crowd scene. Anomalies are detected by conditioning the GAN on normal crowd behavior and identifying deviations.

4. Variation Auto encoders (VAEs)

VAEs are a type of auto encoder that learns a probabilistic latent space representation of the input data, useful for anomaly detection.

VAE-based Anomaly Detection: A VAE is trained on normal crowd data, learning a distribution of latent variables. Anomalies are detected by computing the likelihood of observed data under this learned distribution. Low likelihood indicates an anomaly.

5. Hybrid Models

Combining multiple deep learning models to leverage their strengths for more robust anomaly detection.

Hybrid CNN-RNN Models: These models use CNNs for spatial feature extraction and RNNs for temporal sequence learning, providing a comprehensive approach for detecting anomalies in both spatial and temporal dimensions.

Ensemble Methods: Combining the outputs of multiple models (e.g., CNN, RNN, GAN) to improve the reliability of anomaly detection.

6. Attention Mechanisms

Attention mechanisms help models focus on the most relevant parts of the input data, improving the detection of subtle anomalies.

Attention-Based RNNs: Incorporating attention mechanisms into RNNs allows the model to focus on critical moments in the crowd behavior sequence, enhancing the detection of anomalies.

Self-Attention Networks: These networks, used in models like Transformers, can capture long-range

dependencies and interactions in the data, making them effective for identifying anomalies in complex crowd behaviors.

7. Spatial-Temporal Graph Convolutional Networks (ST-GCN)

These networks model crowd dynamics by representing the crowd as a graph where nodes represent individuals and edges represent interactions.

Graph-Based Anomaly Detection: ST-GCNs analyze spatial-temporal data to identify anomalies in the

Retail Analytics: Detecting abnormal customer behaviors to prevent theft or identify unusual patterns.

Each of these deep learning techniques offers unique advantages for anomaly detection in crowd analysis, and often a combination of these methods is employed to achieve more accurate and robust detection.

patterns of interactions and movements within the crowd.

Applications

Public Safety and Surveillance: Detecting suspicious behaviors, overcrowding, and emergencies.

Transportation Hubs: Monitoring unusual activities in airports, train stations, and bus terminals.

Event Management: Ensuring crowd safety by identifying potential hazards during large gatherings.

Strengths:

1. **Comprehensive Coverage:** The paper covers a wide range of deep learning techniques and their applications in anomaly detection, providing a solid overview of the current state of the field. The inclusion of different models, such as CNNs, RNNs,

applications in public safety, healthcare, transportation, and event management.

3. **Focus on Ethical Considerations:** The discussion on the use of Unmanned Aerial Vehicles (UAVs) for crowd monitoring and the emphasis on policy and ethical considerations is commendable.

Paper	Proposed techniques	Limitations	Research gap
[1]	Convolutional Neural Networks (CNNs)	High computational cost	Optimizing for real-time processing
[2]	YOLO-based Crowd Detection	Struggles with low-light conditions and low-quality images	Enhancing detection accuracy under varying lighting conditions and image quality variations
[3]	Optical flow and generative adversarial networks (GANs)	Scale and illumination variations, object tracking speed, background clutter, object variety, and inconsistencies in camera position and angle.	The need for advancements in unsupervised real-time anomaly detection applications, particularly in crowded scenes.
[4]	Unmanned Aerial Vehicles (UAVs)	Privacy concerns and ethical implications of drone surveillance.	Further studies should explore the integration of drones with other emerging technologies, such as AI, blockchain, and IoT, to enhance crowd monitoring capabilities
[5]	Faster RCNN with Deep Reinforcement Learning Model (IVADC-FDRL).	Need for extensive computational resources due to the complexity of deep learning models like Faster RCNN and DQL.	The need to design more effective models to achieve maximum accuracy and efficiency in real-world surveillance scenarios.
[6]	CNNs with Long Short-Term Memory (LSTM)	Handling overlapping objects and the need for high computational resources due to processing multiple data streams.	The challenge of recognizing abnormal behaviors in crowded scenarios with multiple individuals, which current models struggle to address effectively.
[7]	Deep Learning Architecture	Difficulty in gathering genuine information on abnormal behavior in dense crowds	Further investigation into the integration of multiple techniques, such as combining CNN and K-Means algorithms, to enhance the accuracy and efficiency of abnormal behavior detection in crowds.
[8]	RCNN-based deep learning framework	The presence of false positives and false negatives in the detection results.	Improving the efficiency and accuracy of object detection methods compared to existing techniques.
[9]	Gated Recurrent Unit (GRU), CNN, LSTM architectures	Data Scarcity, Sensitivity to Environmental Changes	Optimization difficulties in model tuning and limitations in motion detection on edge devices.
[10]	Temporal Shift Module (TSM) and ResNet50 Architecture	Specific hardware configurations or computational resources that could be prohibitive for widespread adoption in elevator systems.	Exploration of real-time implementation challenges and considerations for deploying the proposed abnormal behavior detection model

GANs, and hybrid approaches.

2. **Real-World Applications:** The paper effectively connects theoretical models with practical

4. **Identification of Research Gaps:** The paper does a good job of identifying the limitations and challenges of existing models.

Weaknesses:

1. Lack of Detailed Case Studies: While the paper provides a broad overview of different techniques, it lacks detailed case studies or examples of where these methods have been successfully implemented in real-world scenarios.
2. Limited Discussion on Computational Costs: Although the paper mentions the high computational cost of certain models, it does not explore into possible solutions or optimizations to address this issue. A more detailed exploration of how to reduce computational demands would be beneficial, especially for large-scale implementations.
3. Ethical Considerations Could Be Expanded: Although the paper touches on ethical issues related to drone surveillance, it could benefit from a more in-depth discussion on privacy concerns, especially in the context of public surveillance.
4. Minor Language Issues: The paper has some minor language and grammatical issues that could benefit from careful proofreading.

CONCLUSION

This paper offers a valuable survey of the current technologies used for detecting abnormal behaviors in crowds, with a strong focus on deep learning methods. It successfully highlights the potential applications of these technologies in improving public safety and security. However, the paper could be enhanced by including more practical examples, addressing computational challenges in greater detail, and expanding the discussion on ethical considerations. Despite these areas for improvement, the paper makes a significant contribution to the field and is a useful resource for both researchers and practitioners.

REFERENCES

- [1] Learning, From Cctv Footages Using Deep. "Crowd Counting And Anomaly Detection From Cctv Footages Using Deep Learning Augmented With Cellular Automata." *Journal of Theoretical and Applied Information Technology* 101.13 (2023).
- [2] Gündüz, Mehmet Şirin, and Gültekin Işık. "A new YOLO-based method for real-time crowd detection from video and performance analysis of YOLO models." *Journal of Real-Time Image Processing* 20.1 (2023): 5.
- [3] Aldayri, Amnah, and Waleed Albattah. "Taxonomy of anomaly detection techniques in crowd scenes." *Sensors* 22.16 (2022): 6080.
- [4] Husman, Muhammad Afif, et al. "Unmanned aerial vehicles for crowd monitoring and analysis." *Electronics* 10.23 (2021): 2974.
- [5] Mansour, Romany F., et al. "Intelligent video anomaly detection and classification using faster RCNN with deep reinforcement learning model." *Image and Vision Computing* 112 (2021): 104229.
- [6] Kuppusamy, P., and V. C. Bharathi. "Human abnormal behavior detection using CNNs in crowded and uncrowded surveillance—A survey." *Measurement: Sensors* 24 (2022): 100510.
- [7] Sonkar, Riddhi, et al. "Crowd abnormal behaviour detection using deep learning." *ITM Web of Conferences*. Vol. 32. EDP Sciences, 2020.
- [8] Dike, Happiness Ugochi, and Yimin Zhou. "A robust quadruplet and faster region-based CNN for UAV video-based multiple object tracking in crowded environment." *Electronics* 10.7 (2021): 795.
- [9] Wastupranata, Leonard Matheus, Seong G. Kong, and Lipo Wang. "Deep Learning for Abnormal Human Behavior Detection in Surveillance Videos—A Survey." *Electronics* 13.13 (2024): 2579.
- [10] Lei, Jingsheng, et al. "A Model for Detecting Abnormal Elevator Passenger Behavior Based on Video Classification." *Electronics* 13.13 (2024): 2472.