# CROWD ANOMALY DETECTION USING CNN/RNN

Dr B Indira, Assistant Professor, Department of MCA, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad, Telangana State, India
**Shivapriya** MCA Student, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad, Telangana state, Hyderabad

## ABSTRACT

As closed-circuit television (CCTV) surveillance systems have expanded in public settings, crowd anomaly detection has grown in importance as a part of intelligent video surveillance systems. Selecting the captured event demands labour and constant attention, which is difficult to do manually. Crowd monitoring requires more potent anomaly detection techniques. Several modern approaches have been successful in identifying a wide range of aberrant crowd behaviours thanks to the use of geographical and temporal data retrieved from videos. Reducing the model complexity that increases computational and memory demands is an important factor to take into account when it comes to the fast detection of anomalies. In this paper, a low-cost computer method to detect crowd irregularities is proposed.. The suggested solution uses CNN/RNN to generate high recognition accuracy at inexpensive processing cost, to do away with the pricey optical computations. we will conduct experiments Using publicly available datasets in an effort to improve the detection accuracy.

## I. INTRODUCTION

The World Health Organisation (WHO) defines significant gathering events as any occurrence, whether planned or unplanned, that attracts a sizable crowd and places strain on the neighbourhood, city, or nation that is hosting the event. For local organisers working to ensure the event's effective management, the crowd's diverse managing a diverse population that differs in colour, age, language, and culture presents difficult administrative tasks. Administrative authorities are more focused on comprehending the dynamics of crowds, which explain what could be dangerous in massive gatherings. A monitoring programme that quickly discovers anomalies is known as an anomaly detection system. and takes into account any indications of unusual or irregular behaviour.. Due to the extensive use of video surveillance techniques, it is now difficult, time-consuming, and ineffective to manually evaluate the massive amounts of data collected from CCTV cameras. To determine if the captured behaviours are normal or aberrant, it requires workforce and constant attention. Anomalies in crowd scenarios must be correctly identified and detected in order to that, surveillance systems must have an automatic anomaly detection functionality. Rapid and automatic detection of anomalous behaviours in crowded contexts is crucial for enhancing safety, reducing hazards, and ensuring speedy reaction. Anomaly detection in surveillance systems is essential for ensuring safety and, in some cases, the potential for disaster prevention.

There is a growing need for surveillance video monitoring of public scenes due to the numerous new difficulties in public administration, security, and safety. At first glance, it appears to be a simple task for a human to watch the feed from security cameras, extract important and useful information from behavioural patterns, identify anomalous behaviours, and offer an immediate response. However, it is challenging for a person to keep track of multiple signals at oncebecause of severe constraints in human behaviour. It takes a lot of time and requires a lot of resources, including personnel and space. To do this, an automatic detection technique is necessary. The detection of aberrant events is one of the subdomains of behaviour understanding from surveillance cameras. The process of anomaly identification in surveillance cameras can encounter a variety of issues. (1) Because anomalous occurrences are uncommon, it is challenging to locate large databases of them.

The learning process could be hampered by the lack of samples. (2) In general, anything that deviates from a predetermined pattern (or rule) is referred to as a "anomaly". Thus, we are unable to create a model specifically for aberrant events. (3) Depending on the circumstance, a behaviour may be normal or aberrant. It implies that, under some circumstances, even a global abnormal event (GAE)—such as shooting in a gun club—can be a common occurrence.. While "shooting" is typically regarded as aberrant, it appears normal in a shooting club.And, certain behaviour would constitute an anomaly in a particular place and circumstance known as a local abnormal event (LAE), even though it is not always intrinsically abnormal.

Detecting anomalies is a crucial and well-known split of learning techniques into supervised, unsupervised, and semi-supervised procedures. There are two alternative methods of supervised learning, depending on whether the model is trained by a single category or by all of the categories that are currently present. Only normal (or abnormal) events are utilised to train the model in single model learning, whereas both normal and abnormal events must be learned in multi-model learning. By learning a threshold for normalcy definition, a multidimensional model of typical occurrences inside the feature space, and rules for model definition, anomalous events are discriminated from normal ones in the single model learning process. Each class in the multi-model learning technique will receive independent or dependent training, which is especially helpful when there are many groups of anomalies.

## II. LITERATURE REVIEW

Yang et al. [1] presented a two-channel system architecture. The feature channels that make up the scheme are made using the original video's structure. To guarantee that the channels constantly produce two anomaly scores and high level feature representations, two hybrid deep learning architectures are combined. The design combines Deep Belief Networks, a Stacked Denoising Auto-encoder, and a Plane-based One Class SVM. Anomalous Event Detection is a fusion method that combines the anomaly scores and finally aids in crowd anomaly detection.

B. Pradeepa et al. [2] presented system that combines streak flow approaches with the Latent Dirichlet Allocation, sometimes known as LDA. The suggested methodology calls for building blocks out of divided frames that precisely reflect both spatial and temporal scene changes. The optical flow algorithm is used to estimate the direction and motions of a person among a crowd. The motion of the crowd is also influenced by the streak line and potential functions.

Liu et al.[3] specified that, the Gaussian averaging models' sluggish convergence rate is a drawback when used for object recognition. Hence, an improved attenuation technique based on learning rate was introduced. Following analysis of the foreground data set, a predictive neural network is trained using the foreground data set. The discrepancy between the real and predictive frames is measured in order to ascertain the level of irregularity. By altering the threshold in response to specific circumstances, the anomalous behavior of the crowd is discovered.

Guo et al. [4] has put forth a system that may perform processing and incorporate a robot that analyses embedded movies. K-means algorithms are subjected to improvisation. In order to effectively detect anomaly in congested areas, the system advises adopting the MKSM methodology.

Direkoglu et al. [5] offered a technique to understand the location of moving objects where optical flow vectors generate MII which are static image templates. The MII(motion information image) is utilized to train a CNN that is ultimately employed for the detection of anomalies in crowd behavior. Using MII makes it simpler to identify anomalous behavior because it allows you to see how the crowd is moving.

Kulshrestha et al. [6] suggested a surveillance system dubbed SmartISS that uses real-time MAC id tracking and monitoring to identify, track, and monitor a person's wireless device(s) in real time.. The PSUs, or portable trackers,

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

accept user probe requests and their locations without the users' active participation. These PSUs employ a noisy server to store the acquired traces and are made up of a smartphone, a jetson-TK1, and a computer. In turn, the cloud server aids in locating questionable individuals. The suggested LLTR algorithm chooses the most advantageous number of PSUs to dynamically locate those people.

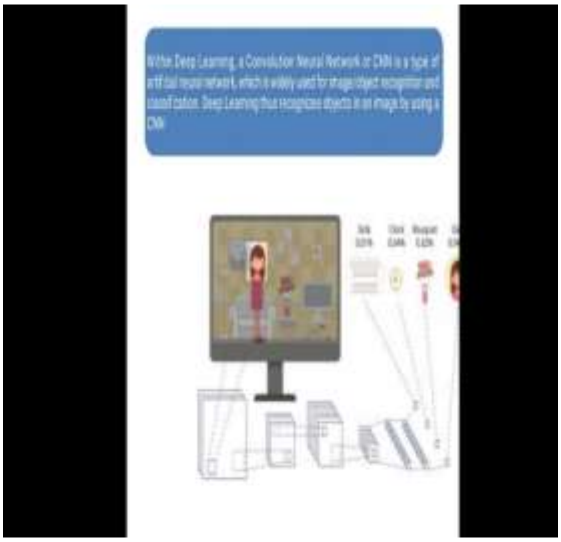
Kong et al. [7] proposed a method that makes use of an LSTM network, where traffic prediction is done to evaluate the disparity between real and projected flows, then refined to produce anomalous characteristics. The abnormal zones are then discovered using OCSVM (One-Class Support Vector Machine).dependent or independent training

## III. METHODOLOGY

### ALGORITHMS

CNN: CNNs (Convolutional Neural Networks) are DL algorithms commonly used for image analysis and recognition. A wide range of computer vision applications have been successfully implemented using CNNs, including image classification, object detection, segmentation, and others.

The figure below shows the functionality of different layers ,



Fig(1) CNN performs frame splitting and focusing by using these five layers.

RNN: RNN stands for Recurrent Neural Network. Unlike feedforward neural networks, which process input data independently, RNNs maintain an internal memory to capture and utilize information from previous time steps. They are specialized in processing sequential data or data with temporal dependencies.

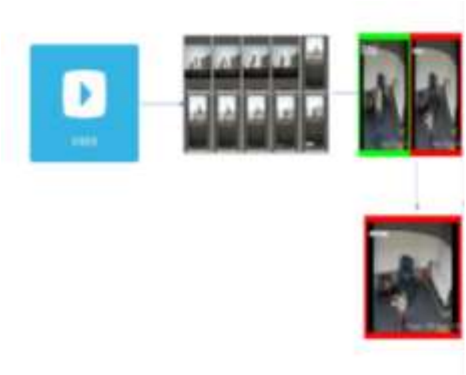
An important characteristic of RNNs is their ability to process sequential data of varying lengths. They preserve a recollection of previous inputs in addition to processing input sequences of various lengths. As a result, RNNs are effective at a variety of tasks, including language modelling, speech recognition, sentiment analysis, time series prediction, and machine translation.

As a result, RNN is mainly used for the Temporal modeling and also because it saves the data that is given to it every time and uses it for the future references. This helps the model to detect the different types of anomalies in the given input.

The video dataset used in this project was collected from Kaggle and social-media platforms. Kaggle is a well-known online platform where it offers various datasets across variety of subjects.

The video data underwent preprocessing technique to extract individual frames and convert them into a suitable format for deep learning models. Here the unwanted artifacts like noise, blank images, blur images have been eliminated like noise , blank images, blur images have been eliminated

ISSN: 2278-4632
Vol-13, Issue-11, November 2023



Fig(2)

Figure (2) shows the process of anomaly detection.

An inserted video is divided into different frames using CNN, the frames are then classified as normal or abnormal, and only the abnormal frames are displayed.

## IV RESULTS

The video will be divided into different frames from which the anomality recognised frames will be separated and stored.

The proposed system utilizes CNN and RNN to detect anomaly activities. The input video undergoes CNN that process the video and split the video into frames and thoroughly analyze each extracted frame. The results are computed as follows:

Firstly the video is splitted into frames using CNN layers. Then using RNN from all the splitted frames the anomality detected frames are seperated and saved. This makes the detection of anomality easy, fast and accurate.

This project has attained the accuracy of 92%.

## V CONCLUSION

In order to identify anomalous behaviour, this work offers a unique structure that combines CNN and RNN. We encountered a number of restrictions when putting this idea into practise. The dataset we used include a variety of individuals, speeds, and lighting conditions. For instance, the video contained various oddities, although in other videos, no one could be seen. In addition, we must address yet another dataset restriction. It may just take one or two seconds for the odd events to occur, and even in 10-second recordings, more than 80% of the time demonstrates that the behaviour is normal. Our suggested method performs better than previous ways despite the constraints indicated. The same background and objects are used for both extraordinary and regular events. We used ResNet50, one of the most popular CNNs, to incorporate the most crucial features from each input frame of video. A ConvLSTM structure is then applied to each ResNet output in order to examine the aberrant event over a number of frames. In order to determine how the model correctly identifies the appropriate category for each input video, we employed classifiers for each dataset.

## VI FUTURE ENHANCEMENT

In order to detect anomalies in busy regions, it is still more important than ever to perform better and be more accurate. There are still numerous issues that need more research, despite the fact that there have been many studies on recognising abnormal human behaviours. Crowd abnormal behaviour identification should be more precise and resistant to a variety of circumstances in vast and diverse crowds. Drones and satellites that use advanced technology to observe the crowd will contribute more insightful information.

## VII REFERENCES

[1] M. Yang, S. Rajasegarar,"A comparative study between single and multi-frame anomaly detection and localization in recorded video streams," J. Vis. Commun. Image Represent., vol. 79, p. 103232, 2021, pp. 1–8, Doi: 10.1209/IkCNN.2019.8852356, 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary. [2] B. Pradeepa, The 2019 International Conference on Wireless Communications, Signal Processing, and Networking (WiSPNET), " Laser-based algorithms meeting privacy in

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

surveillance: A survey,"Chennai, India, 2019, pp. 363-369, Doi:

10.1209/WiSPNETT45539.2019.9032745.

[3] Y. Liu, K. Hao, X. Tang and T. Wang, " Using a predictive neural network, abnormal crowd behaviour can be detected, 2019, pp. 222-225, Doi: 10.1109/ICAIICA.2019.8874488.

[4] Guo Differential privacy preservation in deep learning: Issues, Possibilities, and Solutions. 2019's IEEE Access, volume 7, pages 48 901-48 911.

[5] C. Direkoglu, "Classification of histopathological biopsy images using ensemble of deep learning networks," in 2019 Annual International Conference on Computer Science and Software Engineering (CASCON), Markham, Ontario, Canada, pp. 92–99.

[6] T. Kulshrestha Transactions on Mobile Computing, "Real-Time Crowd Monitoring Using Seamless Indoor-Outdoor Localization," vol. 19, no. 3, 1 March 2020, pp. 664-679; doi: 10.1109/TMC.2019.2897561.

[7] . Kong,

"HUAD: Based on Spatio-Temporal Data," IEEE Access, vol. 8, 2020, pp. 26573-26582, Doi: 10.1109/ACCESS.2020.2971341.