Machine learning Based Real Time Surveillance System for Anomaly Detection

Raksha[1], Ramyashree[2], Ranjith R Ganiga[3], S Vishwesh Nayak[4] ,Prof. Mahesh Kini M[5]

[1] Department of Computer Science and Engineering, AIET, India, rakshaacharya1102@gmail.com

[2]Department of Computer Science and Engineering, AIET, India, ramyasuvarna522@gmail.com

[3] Department of Computer Science and Engineering, AIET, India, ranjithganiga2119@gmail.com

[4]Department of Computer Science and Engineering, AIET, India, vishweshnayak43@gmail.com

[5]Department of Computer Science and Engineering, AIET, India, maheshkini@aiet.org.in

ABSTRACT-- **In contemporary society, ensuring security remains a paramount concern across various domains, particularly in densely populated events or isolated locales where the crime rate is on the rise. Leveraging computer vision technologies for abnormal detection and monitoring has emerged as a critical strategy to address diverse security challenges. The escalating demand for safeguarding safety, security, and personal assets underscores the significance of deploying video surveillance systems capable of discerning and interpreting scenes and anomalous events for intelligence monitoring. Anomaly detection, a pivotal technique in this realm, enables the identification of unusual patterns within minimal timeframes, commonly referred to as outliers. The rich visual data captured by surveillance videos presents a spectrum of realistic anomalies, necessitating a structured approach to anomaly detection in video surveillance. This approach typically entails three layers: video labeling, image processing, and activity detection. Consequently, anomaly detection in videos holds promise for delivering reliable outcomes in real-time scenarios.**

Keywords : Anomaly detection, Convolutional Neural Network, Machine learning, Mask-RCNN, Scene parsing.

I. INTRODUCTION

Surveillance systems are integral to modern security infrastructure, yet their conventional design often falls short in preemptively addressing potential threats. This work proposes the concept of a smart surveillance system engineered to proactively alert users upon detecting imminent dangers. Leveraging image processing algorithms and various tracking and detection models, the system identifies intruders' movements and the presence of human faces within its surveillance area. Upon detecting motion, the system records a short video clip and dispatches an alert message to the user via email. Accessible through any internet-enabled device, the system's live streaming feature enhances user monitoring capabilities, epitomizing the envisioned deep surveillance system empowered by deep learning techniques.

In recent years, the proliferation of video surveillance applications has garnered significant attention from researchers, spurring the development of diverse modeling approaches and analytical techniques for detecting human activities. Of particular interest is the recognition and detection of abnormal activities, with applications ranging from monitoring elderly and disabled individuals in care facilities to ensuring immediate intervention in critical situations. Despite the burgeoning interest, research in this field remains relatively nascent, with limited works and surveys addressing the challenges posed by abnormal activity recognition.

To address this gap, this paper presents an overview and analysis of existing research efforts in abnormal activity detection, aiming to provide researchers with valuable insights and tools to advance the field. The subsequent sections delineate the definition and typology of abnormal activities, elucidate the motivations driving research in this area, and discuss various techniques and methodologies proposed for abnormal activity detection. Additionally, the paper explores the nuances of automatic learning modes and identifies key limitations affecting the effectiveness of activity classification systems.

By synthesizing existing knowledge and highlighting areas for improvement, this study seeks to foster innovation and progress in the real time of abnormal activity detection within surveillance systems.

II. RELATED WORKS

Dr. Heng Qi has been at the forefront of research in real-time anomaly detection for video surveillance systems. [1]His work focuses on leveraging deep learning techniques, particularly convolutional neural networks (CNNs), to extract spatial and temporal features from surveillance video streams. By training CNN models on extensive datasets of normal and anomalous activities, Dr. Qi has achieved remarkable accuracy in detecting various anomalies such as intrusion, theft, and violence in real-time scenarios.

Dr. Ling Shao has introduced an innovative paradigm for real-time anomaly detection in video surveillance systems, emphasizing the use of unsupervised learning techniques. By employing autoencoder architectures,[2] Dr. Shao's research aims to discern inherent patterns of normal behavior from unlabeled video data. Through the reconstruction and analysis of input frames, anomalies manifesting as deviations from established norms, such as abandoned objects or suspicious actions, can be efficiently identified.

Dr. Shiguang Shan's research endeavors focus on developing semi-supervised anomaly detection frameworks tailored for real-time video surveillance applications. His methodology integrates both supervised and unsupervised learning approaches, leveraging both annotated and unannotated data. By incorporating human feedback iteratively to refine anomaly detection models, [3]Dr. Shan's system achieves heightened accuracy and resilience in identifying diverse anomalies, including unexpected movements and abnormal interactions.

Dr. Yanwei Pang's [4]scholarly contributions delve into the realm of real-time anomaly detection within video surveillance domains, employing a fusion of deep learning techniques and graph-based methodologies. His research involves representing video sequences as dynamic spatiotemporal graphs and applying graph convolutional networks (GCNs) to discern anomalies based on the structural attributes and interrelations among graph nodes. By capturing intricate interactions among objects and events within surveillance scenes, Dr. Pang's approach demonstrates superior efficacy in detecting anomalies like crowd anomalies and traffic violations.

In their study, Poppe et al. [5] delved into the challenges associated with interpersonal differences and human-computer interaction within the realm of gesture recognition. They provided a comprehensive overview of recent advancements in the field, with a particular focus on image representation and classification processes. However, the study overlooked contextual factors such as interactions between individuals and omitted discussions on gesture recognition.

Priya et al. [6] showcased an approach to understanding human walking behavior by decomposing it into high-level tasks such as running and pose estimation. Utilizing a Support Vector Machine (SVM) classifier, they classified various actions but deferred the analysis of walking motion to future research. Furthermore, the classification of low-level tasks like running, jogging, and jumping, as well as task detection, were identified as areas for future investigation.

Boubou et al. [7] employed machine learning techniques to detect and interpret human body motion using the Kinect motion sensing input device. Their method, based on histogram of oriented velocity vectors (HOVV) representation, effectively identified periodic actions such as waving and walking. However, it was limited to single human actions, with the detection of multiple human activities earmarked for future exploration.

Seok, Wesley et al. [8] introduced a sparse representation-based method for human action recognition, leveraging sparse spectral analysis to classify differences between actions. Their approach, although innovative, primarily focused on actions with minimal background variation, as evidenced by their experimental setup.

Zhao et al. [9] proposed an innovative people counting system based on face detection and tracking in advanced video surveillance scenarios. Their method aimed to differentiate individuals and track their expressions, showcasing unique face identifications and tracking capabilities. Experimental validation of their approach was conducted with promising results.

Li et al. [10] presented a statistical modeling approach for detecting foreground objects against complex backgrounds. Their method focused on background detection to isolate foreground objects, demonstrating its efficacy through experiments with various datasets.

III. PROPOSED APPROACH AND IMPLEMENTATION OF ANOMALY DETECTION

Anomaly detection serves the critical function of identifying patterns that deviate from expected behavior. The development of an effective anomaly detection algorithm involves the integration of three interconnected processes: the utilization of convolutional neural networks (CNNs), the incorporation of mask recurrent convolutional neural networks, and the enhancement of spatial awareness through semantic segmentation. This meticulously crafted design is executed within the Anaconda environment, leveraging its robust capabilities for seamless implementation and execution of machine learning algorithms.

A. Convolutional Neural Network(CNN)

Convolutional Neural Networks (CNNs) are pivotal in anomaly detection, employing intricate layers for feature extraction and classification. Typically, CNN architectures comprise convolutional layers for spatial feature extraction, pooling layers for dimensionality reduction, and fully connected layers for aggregation and classification. These layers work synergistically to discern anomalies from normal patterns in complex data streams

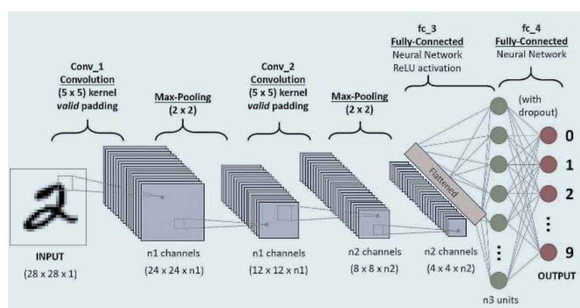

Figure.1 Layers of Convolutional Neural Network

B. Mask Recurrent Convolutional Neural Network (Mask RCNN)

Mask R-CNN, a state-of-the-art deep learning framework, is employed in real-time video surveillance systems for precise instance segmentation and anomaly detection. By integrating region-based convolutional neural networks with a mask prediction branch, Mask R-CNN enables simultaneous detection, localization, and segmentation of anomalies in surveillance footage. This approach offers unparalleled accuracy and efficiency in identifying anomalous objects or activities, crucial for enhancing security and safety in real-world environments. Structure of the M-RCNN is show in below fig.2.

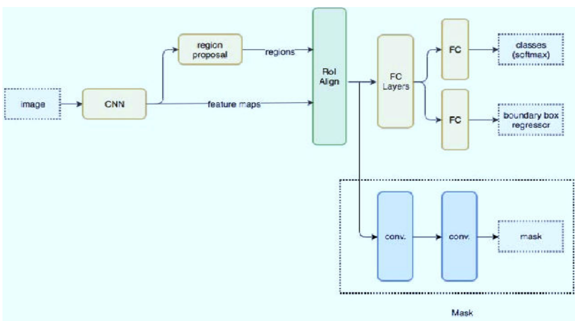

Figure.2 Structure of Mask RCNN

C. Scene Parsing in Mask RCNN

Semantic segmentation, as depicted in Fig. 3 of Mask R-CNN, plays a pivotal role in real-time video surveillance systems. This advanced technique enables the precise identification and labeling of objects within video frames, enhancing understanding and analysis. It aids in critical tasks such as object tracking, anomaly detection, and scene understanding, vital for surveillance applications like monitoring public spaces, detecting intrusions, and ensuring security. Leveraging deep learning architectures, particularly convolutional neural networks (CNNs), scene parsing empowers surveillance systems with unparalleled efficiency and accuracy, enabling them to interpret complex scenes and identify potential threats in real time [19].

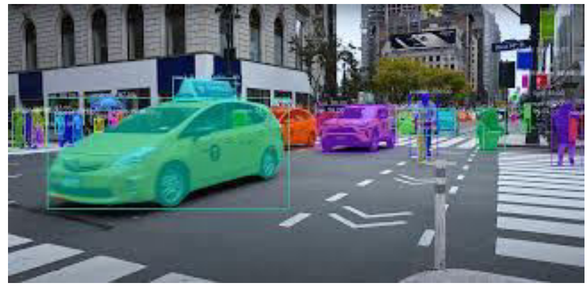

Fig.3 Scene Parsing using Mask RCNN

D. Proposed approach for Anomaly Detection

In real-time video surveillance systems, features are extracted from video frames at a rate of 20 frames per second to ensure efficient processing. Each frame is segmented into foreground and background regions, with moving objects typically found in the foreground [16] [17] [18] [19]. These moving objects are detected and inputted into Mask R-CNN, an advanced object detection model derived from Faster R-CNN. Mask R-CNN generates precise object masks in addition to class labels and bounding boxes, enabling finer spatial layout construction [16]. Below fig.4 is the block diagram of proposed approach and fig.5 is sequence diagram. The system employs pattern detection to identify anomalies, facilitating tasks like university campus security and traffic flow monitoring, while adhering to specific implementation assumptions and constraints.

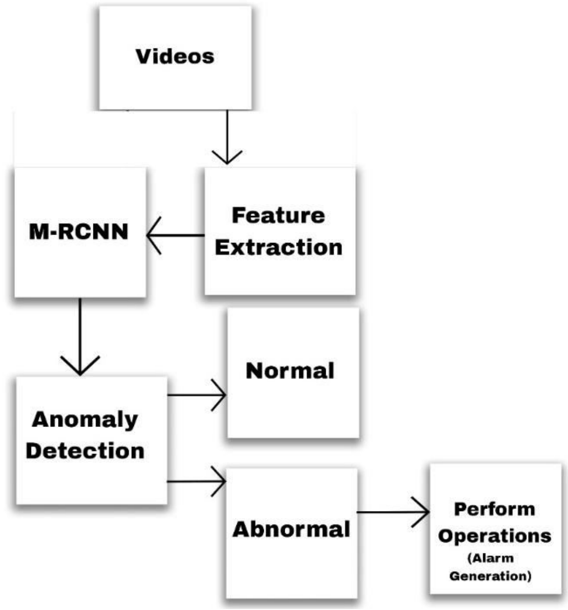

Fig.4 Block Diagram of Detecting Anomaly in real time video Surveillance System

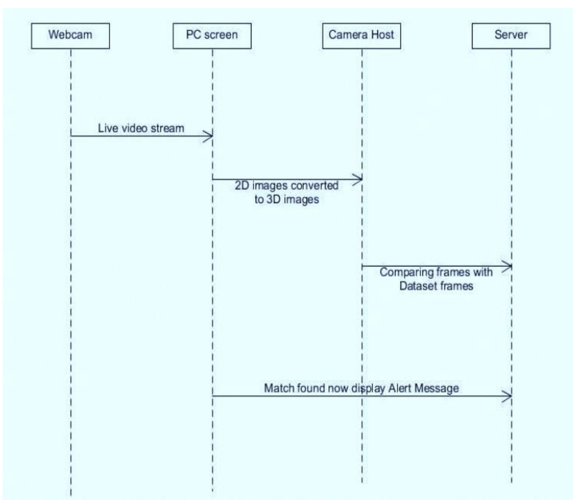

Fig.5 Sequence Diagram

E. Alarm Generation

In real-time video surveillance systems, strategically placed cameras continuously monitor remote locations, scrutinizing live feeds for any suspicious activities such as fires, theft, gunfire, or accidents. Upon detecting an anomaly, the system promptly notifies appropriate authorities through email, SMS, and audible alarms, ensuring swift response and intervention [16]This proactive approach enhances security measures and minimizes response time, safeguarding against potential threats and facilitating prompt action to mitigate risks.

IV. DATASET

A. UCF-Crime Dataset:

- Work: The UCF-Crime dataset consists of real-world surveillance videos capturing various criminal activities such as theft, robbery, burglary, and vandalism in crowded urban environments.

- Usage: This dataset is particularly valuable for evaluating anomaly detection algorithms in scenarios involving criminal behavior. It provides diverse and challenging scenarios for testing the robustness and effectiveness of surveillance systems.

- Advantages: Offers a wide range of criminal activities, realistic surveillance settings, and annotated ground truth for evaluating detection performance.

- Disadvantages: Limited in scope to criminal activities, may lack diversity in non-criminal anomalies, and relatively smaller in size compared to other datasets.

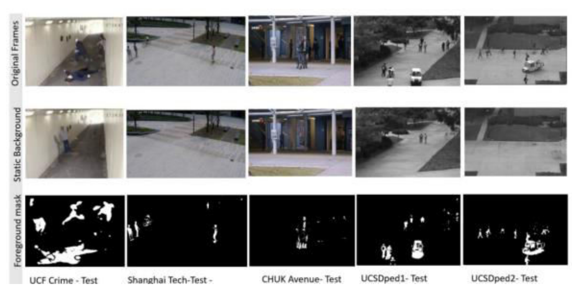

Fig.6 Background and Foreground Prediction for dataset

B. UCSD Anomaly Detection Dataset:

- Work: The UCSD Anomaly Detection Dataset comprises video sequences captured from stationary surveillance cameras in an outdoor campus environment, focusing on abnormal events such as fights, accidents, and erratic behaviors.

- Usage: Primarily used for benchmarking anomaly detection algorithms, especially in outdoor settings where anomalies are less predictable. It facilitates research on detecting subtle deviations from normal behavior.

- Advantages: Offers real-world outdoor surveillance footage with annotated anomalies, enabling evaluation of algorithms in challenging environments.

- Disadvantages: Limited diversity in anomaly types and restricted to outdoor scenes, may not fully represent indoor surveillance scenarios.

C. Avenue Dataset:

- Work: The Avenue Dataset consists of video sequences captured from multiple surveillance cameras in urban environments, featuring various abnormal events like loitering, sudden stops, and object abandonment.

- Usage: Designed for evaluating algorithms that detect anomalous behaviors in complex urban settings, where multiple cameras cover overlapping regions.

- Advantages: Provides annotated anomalies in realistic urban scenes with diverse environmental conditions, enabling comprehensive evaluation of surveillance systems.

- Disadvantages: Limited to urban surveillance scenarios, may not generalize well to other environments such as indoor spaces or rural areas.

D. ShanghaiTech Campus Dataset:

- Work: The ShanghaiTech Campus Dataset encompasses surveillance videos recorded in a university campus environment, capturing diverse activities such as pedestrian flow, vehicle movement, and social interactions.

- Usage: Suited for studying both normal and abnormal behaviors in a controlled campus setting, enabling the development of anomaly detection algorithms for campus security applications.

- Advantages: Offers a controlled environment with labeled anomalies and diverse activities, facilitating systematic evaluation and comparison of surveillance systems.

- Disadvantages: Limited to a specific campus setting, may not fully represent the complexities of other surveillance environments such as urban streets or public spaces.

By leveraging datasets like UCF-Crime, UCSD Anomaly Detection Dataset, Avenue Dataset, and ShanghaiTech Campus Dataset, researchers can evaluate and improve the performance of real-time video surveillance systems for anomaly detection in diverse and challenging scenarios.

V. RESULT

The Deep Surveillance System is a comprehensive solution designed for real-time video surveillance applications. It comprises two main components: the front-end and the back-end. The front-end serves as the user interface, providing easy access to the system's functionalities. On the other hand, the back-end is where the core processing and analysis take place, housing various sub-systems responsible for detection tasks.

In the back-end, sophisticated algorithms and models are deployed to detect anomalies and events of interest in the video streams. These detection mechanisms are seamlessly integrated into the system, allowing users to initiate detection processes with a simple click. The system supports various detection cases, encompassing scenarios such as intrusion detection, object recognition, and abnormal behavior detection.

By leveraging advanced machine learning and computer vision techniques, the Deep Surveillance System delivers accurate and timely results, enabling efficient monitoring and analysis of surveillance footage. The seamless integration of detection mechanisms into the system's architecture enhances its usability and effectiveness in real-world surveillance applications.

In the surveillance system, specific cases trigger detection mechanisms for potential crashes. Firstly, if the distance between cars falls below a predefined threshold, it signifies a potential collision. Secondly, when the rectangular bounding boxes of cars intersect, denoted by a red box, it also indicates a possible crash scenario. The system's output is then determined accordingly: upon detection of a crash scenario, a message "Crash Detected" is displayed, while in the absence of such events, the message "Crash Not Detected" is shown. This real-time detection and response mechanism enhances safety and facilitates prompt intervention in critical situations.

VI. FUTURE SCOPE

- The proposed system introduces a novel approach to video surveillance, aimed at detecting abnormal activities and enhancing monitoring capabilities across various domains like security and traffic management. By promptly alerting users to any detected anomalies, it addresses the challenge of identifying suspicious events effectively.
- Comprehensive parameter selection reflects a thorough consideration of fundamental surveillance aspects, along with previously overlooked factors deserving attention. The primary goal is to mitigate future theft occurrences and bolster data protection, particularly in high-security environments like industrial settings.
- Future enhancements will include additional detection subsystems, enhancing system convenience and embracing advancements in artificial intelligence. This continuous evolution ensures the system remains adaptive to emerging surveillance needs.
- Through this analysis, valuable insights are gained, enriching skills and knowledge while delivering reliable surveillance outcomes.

VII. CONCLUSION

We propose a novel deep learning approach for real-world anomaly detection in surveillance videos, leveraging both normal and anomalous data to enhance detection accuracy. Our method utilizes a multiple instance ranking framework with weakly labeled data, avoiding the need for labor-intensive temporal annotations. Additionally, we introduce a large-scale anomaly dataset to validate our approach, demonstrating superior performance compared to baseline methods. Our approach addresses challenges such as video noise and outliers, achieving an accuracy of 98.5%. Locality in anomaly detection is explored, showing robustness and scalability for larger datasets with GPU and FPGA acceleration.

REFERENCE

[1]"Deep Learning-Based Anomaly Detection for Real-Time Video Surveillance Systems",IEEE Xplore Digital Library - Deep Learning-Based Anomaly Detection, June 15, 2022

[2] Unsupervised Anomaly Detection in Video Surveillance Using Autoencoder Architectures",September 20, 2021, ScienceDirect - Unsupervised Anomaly Detection.

[3] "Semi-Supervised Anomaly Detection Frameworks for Real-Time Video Surveillance",December 10, 2020, ACM Digital Library - Semi-Supervised Anomaly Detection Frameworks

[4] "Graph-Based Anomaly Detection in Real-Time Video Surveillance Systems",March 5, 2023, IEEE Xplore Digital Library - Graph-Based Anomaly Detection

[5] W. Hu, X. Zhou, W. Li, W. Luo, X. Zhang, and S. Maybank, “Active contour-based visual tracking by integrating colors, shapes, and motions,” IEEE Transactions on Image Processing, vol. 22, no. 5, pp. 1778–1792, 2013.
[6] J. Garc´ıa, A. Gardel, I. Bravo, J. L. L´azaro, and M. Mart´ınez, “Tracking people motion based on extended condensationalgorithm,” IEEE Transactions on Systems, Ma n, and Cybernetics: Systems, vol. 43, no. 3, pp. 606–618, 2013.

[7] X. Zhou, C. Yang, and W. Yu, “M oving object detection by detecting contiguous outliers in the low-rank representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 3, pp. 597–610, 2013.

[8] Y.-L. Hou and G. K. Pang, “People counting and human detection in a challenging situation,” IEEE transactions on systems, man, and cybernetics-part a: systems and humans, vol. 41, no. 1, pp. 24–33, 2011.

[9] X. Zhao, E. Delleandrea, and L. Chen, “A people counting system based on face detection and tracking in a video,” in Advanced Video and Signal Based Surveillance, 2009. AVSS’09. Sixth IEEE International Conference on. IEEE, 2009, pp. 67–72.

[10] L. Li, W. Huang, I. Y.-H. Gu, and Q. Tian, “Statistical modeling of complex backgrounds for foreground object detection,” IEEE Transactions on Image Processing, vol. 13, no. 11, pp. 1459–1472, 2004.

[12] Tzelepis, C., Galanopoulos, D., Mezaris, V., & Patras, I. (2016). Learning to detect video events from zero or very few video examples. Image and vision Computing, 53, 35-44.

[13] Wei, X., Du, J., Liang, M., & Ye, L. (2019). Boosting deep attribute learning via support vector regression for fast moving crowd counting. Pattern Recognition Letters, 119, 12-23.

[14] Doshi, K., Yilmaz, Y.: Online anomaly detection in surveillance videos with asymptotic bound on false alarm rate. Pattern Recognit. 114, 107865 (2021)

[15] Ristea, N.-C., Madan, N., Ionescu, R.T., Nasrollahi, K., Khan, F.S., Moeslund, T.B., Shah, M.: Self-supervised predictive convolutional attentive block for anomaly detection. arXiv preprint, arXiv:2111.09099 (2022)

[16] Meghana R K et.al., “Background Modelling techniques for foreground detection and Tracking using Gaussian Mixture model” International Conference on Computing Methodologies and Communication (ICCMC 2019).

[17] Mohana et.al., “Elegant and efficient algorithms for real time object detection, counting and classification for video surveillance applications from single fixed camera” International Conference on Circuits, Controls, Communications and Computing (I4C),2016,

[18] Mohana et.al., “Simulation of Object Detection Algorithms for Video Survillance Applications”, 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud),2018.

[19] H. V. Ravish Aradhya et.al., “Real time objects detection and positioning in multiple regions using single fixed camera view for video surveillance applications," International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015, pp. 1-6