International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 10 Issue 01 | 2025

Future Approach of Real-Time Surveillance Videos for Anomaly Detection with Intelligent Video Surveillance

Vaibhav P. Narkhede[1], Dr. P. M. Jawandhiya[2], Priyanka S. Kharche[3]
[1]Research Scholar, Pankaj Laddhad Inst. of Tech. Buldhana, & MS,
Email: vaibhavnarkhede@hotmail.com
[2] Principal, P.R.Pote College of Engineering & Management, Amravati
Email: pmjawandiya@gmail.com
[3] Assistant Professor, Pankaj Laddhad Inst. of Tech. Buldhana, & MS,

DOI: 10.5281/zenodo.15751583

ABSTRACT

Video surveillance systems (VSS) have been greatly improved by recent developments in computer vision and artificial intelligence. The main goal of a video security system (VSS) is to improve privacy and security by analysing video frames intelligently, which entails supervising or remotely monitoring specified occurrences with cameras and recorders. Dynamic surroundings, the need to increase performance, and changing weather conditions are a few of the problems that affect the accuracy and efficiency of some modern systems that use machine learning and classical image processing. In order to overcome these restrictions, this research suggests an approach that is both efficient and accurate. The system's goal is to aid human observers who might have trouble analysing changes in real-time by including intelligence for video data filtering and interpretation. The suggested method will use video analysis to spot and identify individuals, cars, objects, and suspicious behaviour. Among the many advantages offered by the combination of AI, computer vision, and the IoT are the following: better analysis of audio and video; recognition of movement patterns; tracking of gestures and behaviours; and, finally, a reduction in the need for human resources. Our main purpose is to create a system that can recognise and classify items in real-time from live video streams, specifically focussing on people and cars. The system takes into account the YOLO and SSD models with fusion techniques to maximise anomaly detection accuracy, as well as an alert mechanism, training and validation, an anomaly detection model, feature extraction, and model integration, among other methodologies, to accomplish this project goal.

Index Terms— ***Intelligent video, Activities, Internet of Things, Identifying, Extraction etc.***

1. INTRODUCTION

The expansion of criminal activities endangering lives and property has made the task of maintaining public safety and security more difficult in the modern world. As a result, protecting one's private data, habits, and property has become a top priority for people and communities. By allowing the detection and monitoring of numerous moving objects, surveillance video analysis is vital for spotting unexpected events, aberrant crowd behaviour, and uncommon human activities. In the past, human operators were crucial to traditional video surveillance systems (VSS), watching film and identifying important details that needed to be reported to the public or appropriate authorities in a timely manner. But it's impractical and time-consuming to continuously and uninterruptedly monitor video streams in order to distinguish normal from abnormal events based merely on the history of moving objects. Consequently, sophisticated video surveillance systems (VSS) must be put into place in order to capture a variety of circumstances efficiently and respond quickly [1-3]. The increasing difficulty of managing large crowds and conducting urban surveillance has brought attention to the need for sophisticated technology that can detect abnormal behaviour in such settings. Anomaly detection has been greatly enhanced with the introduction of deep learning, which can now detect incredibly minute changes from the norm. One novel method, Deep Guard, uses advanced deep learning techniques to identify population anomalies. Adapting to the ever-changing dynamics of crowd behaviour, which are impacted by elements such as social events, public gatherings, and urban activities, requires a strong anomaly detection system.

2. LITERATURE REVIEW

According to Muthurasu et al., anomalous behaviour in the contemporary world poses dangers to others. Anything that deviates from the expected, usual, or typical is called an anomaly. The installation of intelligent video surveillance is crucial due to the difficulties of constantly watching public spaces. Research in the area of surveillance video applications has advanced due to the complexity of the subject of detecting anomalous crowd activities. Discovering unusual gatherings, or cases of unusual crowd conduct, is the primary goal of this research. To tackle these challenges, a range of strategies have been utilised, such as histogram representation, optical flow calculation, and algorithms based on deep learning. However, because to obstructions, noise, and congestion, this problem is not adequately addressed. Significant technological improvements were made possible by the development of AI technology. The technology uses a number of methods to distinguish between distinct suspicious actions while it monitors video content in real-time. The difficulty in identifying suspicious from normal human activity stems from the fact that people are so unpredictable. It is usual practice to extract multiple frames at once from a video in order to carry out monitoring. In the framework, you can find two parts. First, the framework calculates the features from each frame of the movie. The classifier then uses these characteristics to decide if the class is normal or panicked. We test the proposed approach on the PETS 2009, MED, and UMN datasets, which are all publicly available. In order to determine how effective the proposed approach is, it is compared to already-established methods. [1]

The number of persons who require the assistance of a wheelchair to move around freely is on the rise, according to Vermander et al. By incorporating technology into these devices, it is possible to objectively evaluate posture and track the functional status of people using wheelchairs at the same time. In this manner, pertinent information regarding the healing process can be shared between the user and the medical staff. Patients can avoid additional musculoskeletal issues and risky circumstances like ulcers or falls by utilising this information to make early adjustments to their rehabilitation programs. As a result, impacted people are encouraged to lead better lives. So, to help find seated irregularities, this study gives a structured review of the current postural diagnosis techniques. Both the monitoring devices and the methodologies utilised for anomaly identification make up this postural diagnosis. The first portion of the analysis focusses on the instruments needed to collect postural data. Here, we'll look at these anomaly detection methods from two angles: first, the conventional, widely-used one, which views anomalies as out-of-the-ordinary positions; second, a more personalised, tailored approach, which views anomalies as deviations from the typical sitting pattern. This allows for an examination of the methods' merits, shortcomings, and potential. In order to help researchers navigate the current state of knowledge in this subject, this overview study primarily aims to synthesise and organise information, detect trends, and provide a thorough grasp of sitting posture diagnosis methods. [2]

The scientific community is increasingly focussing on anomaly identification in video surveillance, according to Mahareek et al. There is a significant need for intelligent systems that can detect anomalies in live video streams automatically. This article provides a comprehensive overview of the many techniques for identifying anomalies in surveillance footage. Statistical modelling and motion analysis are examples of more traditional approaches; deep learning and artificial intelligence are examples of more modern ways. The research also details the potential benefits and drawbacks of each method, as well as their applications in the actual world. Issues with creating effective anomaly detection algorithms for surveillance films are also discussed, and possible directions for future study are suggested. Researchers and practitioners in the field of violent behaviour detection (VioBD) will find it to be an invaluable resource. In addition to shedding light on the current situation of the field, it offers a blueprint for future studies about anomaly detection in surveillance footage. Remembering that the availability and quality of training data greatly affects the success of anomaly identification in surveillance videos is vital for ensuring the best possible performance of the anomaly detection system. Therefore, improving the readability of anomaly detection models and developing trustworthy feature extraction methods should be the focus of future research. According to the study, future research should focus on improving the scalability and effectiveness of anomaly detection systems so that large-scale video data may be utilised in real-world applications. [3]

The worldwide market for drones is expected to reach $5.6 billion in 2020, up from $1.6 billion in 2015, according to Patthe et. al. Drones are becoming more common, but they aren't without their problems. An example of the necessity for appropriate regulation is the 1,000 flights that were disrupted in December 2018

due to drone sightings at Gatwick Airport. In light of these occurrences, it is crucial to guarantee security and mitigate any dangers by effectively detecting, tracking, and identifying anomalous behaviours of small UAVs in complicated situations. The use of data fusion to improve detection performance, particularly when handling heterogeneous data from several sensors, is the main emphasis of this paper's state-of-the-art review of small UAV aberrant behaviour detection. In contexts where numerous drones are in operation, such as Amazon's delivery system, our study provides a systematic review of strategies for detecting anomalous behaviour and highlights the importance of data fusion in solving problems. Furthermore, our research emphasises the importance of promoting standardised performance measures for algorithms that detect anomalous activity. Evaluating behaviour detection in a data fusion system is both challenging and effective, in contrast to the traditional metrics for detection, tracking, and classification, such as accuracy, precision, and MOTA, respectively. [4]
According to Duong et al., the field of anomaly detection in video surveillance is well-established and receiving more and more attention from academics. Intelligent technologies that can automatically identify unusual occurrences in live video streams are in high demand. As a result, numerous strategies for developing a reliable model to guarantee public safety have been put up. Anomaly detection has been the subject of numerous surveys, including those pertaining to networks, financial fraud, human behaviour study, and countless more. Numerous subfields of computer vision have found fruitful uses for deep learning. To be more specific, the suggested strategies primarily employ generative models due to their rapid expansion. The purpose of this article is to survey all the methods employed for video anomaly detection that rely on deep learning. Different methodologies have been established to classify deep learning-based systems according on their goals and learning metrics. Also covered extensively are methods for the vision-based domain's preprocessing and feature engineering. Training and anomalous human behaviour detection benchmark datasets are also detailed in this research. We conclude by outlining some of the most typical problems with video surveillance and suggesting ways forward for both current and future studies.
[5]

3. OBJECTIVES OF PROPOSED SYSTEM

Following are the objectives in which the work will be achieved

- To develop a system capable of real-time object detection and classification using a live camera feed. The primary objective is to accurately identify and classify objects into two main categories: humans and vehicles.
- To design a comprehensive AI-powered surveillance system capable of not only detecting and classifying humans and vehicles in real-time but also identifying suspicious or unusual behaviours.
- To enhance security and situational awareness in the monitored environment.
- To apply methods for reducing the time consumption for training.

4. RESEARCH METHOLOGY

a. YOLOv5:

- Type: One-stage object detection model.
- Approach: Divides the input image into a grid and predicts bounding boxes and object classes for each grid cell.
- Architecture: Composed of three main parts:
    - Backbone: Extracts features from the input image.
    - Neck: Collects and integrates feature maps from different stages of the backbone.
    - Head: Performs the final predictions of bounding boxes and class probabilities.
- Strengths: Known for its speed, accuracy, and compact size, making it suitable for various applications, including resource-constrained devices.

b. SSD (Single Shot MultiBox Detector):

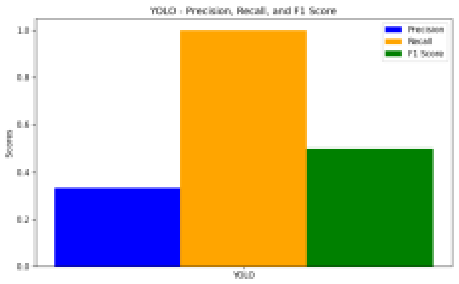
- Type: One-stage object detection model.
- Approach: Directly predicts object classes and their bounding boxes from input images, without a separate region proposal step.
- Key Operations:

- Feature Extraction: Utilizes a base CNN architecture (e.g., VGG, ResNet) to extract feature maps at different scales. These multi-scale feature maps enable the detection of objects of varying sizes.
- Multi-Scale Detection: Employs feature maps from different layers of the network to detect objects of different sizes (earlier layers for smaller objects, deeper layers for larger objects).
- Bounding Boxes and Class Predictions: Generates default boxes (anchors) with multiple aspect ratios at each position of the feature map and predicts both the object class and bounding box offsets in a single step.

- Strengths: Designed to be fast while maintaining good accuracy.

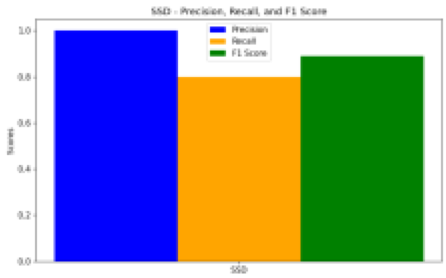
In PHASE 1 we find out the processed output of the YOLO model and in Phase 2 we processed this output as an input to SSD and accuracy with the fusion of the models is high as compared to individual models.
The output of individual model and combined models was given as follows:

YOLO

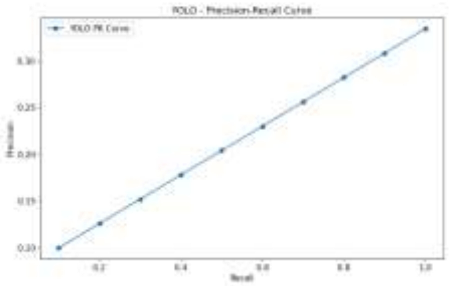


F1 Score

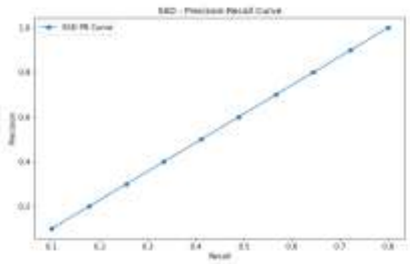
SSD



F1 Score



Precision Recall Curve



Precision Recall Curve

Anomaly Detection

Anomaly Detection

a. YOLO Output Analysis:

- Precision (0.3343): YOLO's precision of 0.3343 indicates that only about 33% of its positive object detections were actually correct. This relatively low value suggests a significant number of false positives, meaning YOLO incorrectly identified many non-objects as objects.
- Recall (1.0): YOLO achieved a perfect recall of 1.0, meaning it successfully detected all the actual objects present in the video. It didn't miss any true objects.

- F1 Score (0.5011): The F1 score, which balances precision and recall, is 0.5011. This moderate score reflects the trade-off between YOLO's perfect ability to find all objects and its tendency to make many incorrect detections.

b. SSD Output Analysis:

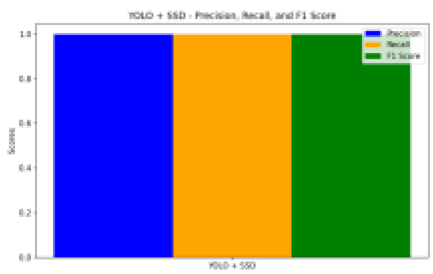
- Precision (1.0): SSD demonstrated perfect precision (1.0), indicating that all of its positive object predictions were accurate. It did not produce any false positives.
- Recall (1.0): SSD also achieved a perfect recall (1.0), meaning it successfully detected all the actual objects present in the video without missing any.
- F1 Score (1.0): With both perfect precision and recall, SSD's F1 score is also a perfect 1.0, signifying flawless object detection performance.

c. Overall Performance Summary:

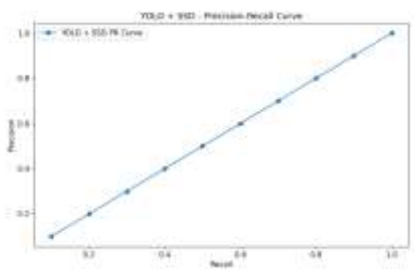
- YOLO: While YOLO successfully detected every object (perfect recall), its low precision (0.3343) indicates a high number of false positive detections. This imbalance results in a moderate F1 score of approximately 0.50.
- SSD: In contrast, SSD exhibited perfect performance across all metrics (precision, recall, and F1 score all at 1.0), indicating that it correctly identified all objects without any false alarms.

5. FUSION OUTPUT

YOLO +SSD



F1 Score



Precesion Recall

Anomaly Detection

6. FUTURE ASPECTS OF INTELLIGENT VIDEO SURVEILLANCE

- Enhanced Predictive Capabilities: Future systems will likely move beyond simple detection and recognition to predict potential security threats and anomalies before they occur. This could involve analyzing patterns of behavior, environmental changes, and contextual information to proactively identify risks.

- Greater Integration and Interoperability: Expect seamless integration of intelligent video surveillance with other security systems (e.g., access control, alarm systems) and broader smart city infrastructure. This will enable more coordinated and comprehensive security responses.
- Improved Privacy Preservation Techniques: As these systems become more sophisticated, there will be a greater focus on developing and implementing advanced privacy-preserving techniques. This could include federated learning, differential privacy, and anonymization methods to analyze video data while safeguarding individual privacy.
- Edge Computing and Decentralized Processing: Future intelligent video surveillance will likely leverage more edge computing to process data locally at the camera or edge device. This will reduce latency, bandwidth requirements, and enhance real-time responsiveness.
- Human-AI Collaboration: The future will likely see more seamless collaboration between human operators and AI-powered surveillance systems. AI will handle routine monitoring and anomaly detection, while human experts will focus on verification, complex analysis, and decision-making.
- Adaptive and Self-Learning Systems: Future systems will become more adaptive and capable of self-learning from new data and evolving environments. This will lead to continuous improvement in accuracy, reduced false alarms, and better generalization across diverse scenarios.
- Expansion into New Domains: The applications of intelligent video surveillance will likely expand beyond traditional security to areas like healthcare (patient monitoring, fall detection), transportation (traffic management, autonomous vehicle safety), retail (customer behavior analysis, loss prevention), and environmental monitoring.
- Increased Focus on Explainability and Trust: As AI plays a more critical role in surveillance, there will be a growing need for explainable AI (XAI) to understand the reasoning behind system decisions, fostering greater trust and accountability.
- Miniaturization and Ubiquitous Deployment: Advances in sensor technology and miniaturization will lead to the deployment of smaller, less obtrusive intelligent cameras in a wider range of environments.
- Standardization and Regulation: As the use of intelligent video surveillance proliferates, we can anticipate the development of more standardized protocols and regulations to address ethical concerns, data privacy, and security standards.

7. CONCLUSION

The ability to identify, monitor, and identify objects through the intelligent analysis of video from strategically placed cameras has a profound effect on privacy and security. To simplify human life and improve accuracy and efficiency, good management and control across all societal sectors are necessary, especially with the rising global deployment of security and monitoring systems to prevent anomalous events. More flexible and reliable video surveillance systems may be built with the use of AI and computer vision, making them safer for people and precious assets with less room for human mistake. As a result, people in every walk of life continue to demand effective security and safety solutions.

8. REFERENCES

[1] Muthurasu N, Rajasekar V, “Unveiling Anomalies in Crowds through Advanced Deep Learning for Unusual Activity Detection”, Journal of Theoretical and Applied Information Technology 15th March 2024. Vol.102. No 5
[2] Patrick Vermander, Aitziber Mancisidor, Itziar Cabanes and Nerea Perez, “Intelligent systems for sitting posture monitoring and anomaly detection: an overview”, journal of NeuroEngineering and Rehabilitation (2024) 21:28
[3] Esraa A. Mahareek, Eman K. ElSayed, Nahed M. ElDesouky, Kamal A. ElDahshan, “Anomaly Detection in Surveillance Videos: Survey”, International Journal of Theoretical and Applied Research, 2024, Vol. 3, No. 1, 328-343
[4] Pierre PATHE, Benjamin Pannetier, Bartheye Olivier, Daniel, “Abnormal Behavior State-of-the-Art for UAV Detection in Complex Environments”, HAL Id:hal-04513213 https://hal.science/hal-04513213 Submitted on 20 Mar 2024
[5] Huu-Thanh Duong, Viet-Tuan Le and Vinh Truong Hoang, “Deep Learning-Based Anomaly Detection in Video Surveillance: A Survey”, Sensors 2023, 23, 5024. https://doi.org/10.3390/s23115024
[6] Neamah, K. A. (2022, January 20). Mathematical Model for Handling Unstable Time Series by Using a Linear Approximation Technique. Webology, 19(1), 2835–2852.
[7] Mustafa, M. (2022). Real-Time Video Anomaly Detection for Smart Surveillance. SSRN Electronic Journal.
[8] Chakole, P. D., Satpute, V. R., & Cheggoju, N. (2022, May 1). Crowd behavior anomaly detection using correlation of optical flow magnitude. Journal of Physics: Conference Series, 2273(1), 012023.

[9] Luque Sánchez, F., Hupont, I., Tabik, S., & Herrera, F. (2020, December). Revisiting crowd behaviour analysis through deep learning: Taxonomy, anomaly detection, crowd emotions, datasets, opportunities and prospects. Information Fusion, 64, 318–335.

[10] Bhuiyan, M. R., Abdullah, J., Hashim, N., & Al Farid, F. (2022, March 29). Video analytics using deep learning for crowd analysis: a review. Multimedia Tools and Applications, 81(19), 27895–27922.

[11] Lalit, R., Purwar, R. K., Verma, S., & Jain, A. (2021, December 14). Crowd abnormality detection in video sequences using supervised convolutional neural network. Multimedia Tools and Applications, 81(4), 5259–5277.

[12] Sonkar, R., Rathod, S., Jadhav, R., & Patil, D. (2020). CROWD ABNORMAL BEHAVIOUR DETECTION USING DEEP LEARNING. ITM Web of Conferences, 32, 03040.

[13] Bamaqa, A., Sedky, M., Bosakowski, T., Bakhtiari Bastaki, B., & Alshammari, N. O. (2022, October). SIMCD: SIMulated crowd data for anomaly detection and prediction. Expert Systems With Applications, 203, 117475.

[14] Sharif, M. H., Jiao, L., & Omlin, C. W. (2023, March 23). Deep Crowd Anomaly Detection by Fusing Reconstruction and Prediction Networks. Electronics, 12(7), 1517.

[15] Jia, D., Zhang, C., & Zhang, B. (2021, January 27). Crowd density classification method based on pixels and texture features. Machine Vision and Applications, 32(2).

[16] Mehmood, A. (2021, April 14). Abnormal Behavior Detection in Uncrowded Videos with Two-Stream 3D Convolutional Neural Networks. Applied Sciences, 11(8), 3523.

[17] Wang, B., & Yang, C. (2022, June 20). Video Anomaly Detection Based on Convolutional Recurrent AutoEncoder. Sensors, 22(12), 4647.

[18] Surucu, O., Gadsden, S. A., & Yawney, J. (2023, July). Condition Monitoring using Machine Learning: A Review of Theory, Applications, and Recent Advances. Expert Systems With Applications, 221, 119738.

[19] Chen, S., & Guo, W. (2023, April 7). Auto- Encoders in Deep Learning—A Review with New Perspectives. Mathematics, 11(8), 1777.

[20] Smart Surveillance: A Review & Survey Through Deep Learning Techniques for Detection & Analysis. (2023, October 5). Journal of Sensor Networks and Data Communications, 3(1).

[21] Sreenu, G., & Saleem Durai, M. A. (2019, June 6). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. Journal of Big Data, 6(1).

[22] Nafea, O., Abdul, W., Muhammad, G., & Alsulaiman, M. (2021, March 18). Sensor-Based Human Activity Recognition with Spatio- Temporal Deep Learning. Sensors, 21(6), 2141.

[23] Nawaz, A., Zhiqiu, H., Senzhang, W., Hussain, Y., Khan, I., & Khan, Z. (2020, April 8). Convolutional LSTM based transportation mode learning from raw GPS trajectories. IET Intelligent Transport Systems, 14(6), 570–577.

[24] Zou, Z., Cheng, Y., Qu, X., Ji, S., Guo, X., & Zhou, P. (2019, November). Attend to count: Crowd counting with adaptive capacity multiscale CNNs. Neurocomputing, 367, 75–83.

[25] Shrestha, S., & Vanneschi, L. (2018, July 18). Improved Fully Convolutional Network with Conditional Random Fields for Building Extraction. Remote Sensing, 10(7), 1135.

[26] Wan, B., Jiang, W., Fang, Y., Luo, Z., & Ding, G. (2021, May 22). Anomaly detection in video sequences: A benchmark and computational model. IET Image Processing, 15(14), 3454– 3465.

[27] Xia, L., Mi, S., Zhang, J., Luo, J., Shen, Z., & Cheng, Y. (2023, May 22). Dual-Stream Feature Extraction Network Based on CNN and Transformer for Building Extraction. Remote Sensing, 15(10), 2689.

[28] Ben Mabrouk, A., & Zagrouba, E. (2018, January). Abnormal behavior recognition for intelligent video surveillance systems: A review. Expert Systems With Applications, 91, 480–491.

[29] Ammar, H., & Cherif, A. (2021, March 10). DeepROD: a deep learning approach for realtime and online detection of a panic behavior in human crowds. Machine Vision and Applications, 32(3). https://doi.org/10.1007/s00138-021-01182-w