OPEN ACCESS JOURNALS
Indiana Journal of Multidisciplinary Research
Abbreviated Tittle- Ind J Multi Res
ISSN *(Online)*- 2583-3820
Journal Homepage Link- https://indianapublications.com/Journals/IJMR
DOI: https://doi.org/10.5281/zenodo.21556878

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article

Volume-06|Issue-04|2026

Suspicious Activity Detection Using Machine Learning

Jyoti Neeli[1], Shwetha S V[2], Rakshitha H M[3], Nuthan[4]

[1,2,3,4]Nitte Meenakshi Institute of Technology (NMIT), Nitte(Deemed to Be University), Department of Computer science & Engineering, Bengaluru, India.

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

Citation
Neeli, J., Shwetha, S. V., Rakshitha, H. M., Nuthan. (2026). Suspicious Activity Detection Using Machine Learning. *Indiana Journal of Multidisciplinary Research, 6*(4), 1-7.

Abstract: Video surveillance has become a critical component in today's world. With developments in advanced systems have been developed as a result of the precision and efficacy of deep learning, machine learning, and artificial intelligence to identify and identify questionable activities while live image monitoring is taking place. Because human behavior is fundamentally unpredictable, it can be difficult to judge whether an action is suspicious or typical. Human activities and behaviors were divided into two categories in this study: suspicious and normal. Suspicious activities include sprinting, boxing, or fighting, whereas normal activities include sitting, strolling, jogging, and waving their hands. Convolutional neural networks are used to achieve this classification (CNNs). High-level features are first extracted from the photos by the CNN. The convolutional network's classification is then applied, and the final prediction is informed by the outcomes of the final pooling layer.

Keywords: Video Surveillance, Suspicious activity detection, Human activity recognition, Convolutional Neural Network (CNN), Deep learning, Intelligent security systems.

INTRODUCTION

Although security cameras are commonly utilized, crime rates continue to escalate in today's society. To effectively identify suspicious behavior and enable swift responses, a system that minimizes identification time must be developed. One viable approach is to process surveillance footage by breaking it down into individual frames and then analyzing these images.

While there are various machine learning techniques for image analysis, their accuracy often declines as the dataset increases in size. Deep learning algorithms are investigating methods to enhance performance despite this challenge. The installation of CCTV cameras for constant surveillance has become standard practice across numerous industries and sectors. In densely populated, developed nations, individuals are frequently captured on camera multiple times each day. Consequently, enormous volumes of video data are generated and stored for specific periods. It is quite difficult to manually monitor such surveillance footage to spot suspicious activity because it takes a committed staff and ongoing attention to detail. In public places including parking lots, prisons, military installations, places of worship, borders, and train stations, automated video surveillance systems provide a useful answer. These devices have the ability to stop things like congestion and physical fights, theft, vandalism, carrying explosives or weapons, and other destructive actions. Automated video monitoring is also crucial for improving ATM and bank security. Bank surveillance systems can prevent armed robberies and heists.

Automated monitoring improves security for ATMs, which are frequently the target of theft. When it comes to spotting disruptive student behavior on campuses, including bullying or fighting, security cameras can be extremely helpful. Additionally, they can help to make anti-theft policies in schools stronger. Automated surveillance systems are utilized in test rooms to find suspicious activity, such as cheating or unauthorized behavior. Security cameras are being used more and more by industries, shopping malls, and small companies to prevent armed robberies, stealing, and other criminal activity. Additionally, useful in warehouses for inventory tracking and staff detection, these cameras detect misbehavior such as theft or bribery, as well as maintaining general operational security.

Surveillance systems are used to keep an eye on platforms, routes, highways, tunnels, and parking lots in public transportation hubs including train and bus stations. Malicious activity, such **as** terrorists leaving unsecured backpacks containing explosives, may target such sites. Automated security cameras can detect abandoned objects and notify **authorities**, allowing them to take preventative action to protect infrastructure and passengers.

Video surveillance can help with patient and senior citizen home observation in healthcare settings. These technologies enable caretakers to react quickly by **recognizing** anomalous behavior, such as fainting, vomiting, or other anomalies. Because surveillance technology is used in so many different fields, it is crucial

***Corresponding Author:** Jyoti Neeli

to develop sophisticated methods for spotting questionable activity in video footage.

LITERATURE REVIEW

Real-time video analytics has emerged as a promising approach for enhancing surveillance systems by integrating machine learning models with live-stream monitoring for anomaly detection. Gayathri *et al.* [3] developed a monitoring system utilizing UAVs to provide wide-area coverage in security surveillance, leveraging deep learning algorithms to identify abnormalities within aerial footage. Similarly, Bora and Rokade [4] proposed a convolutional neural network (CNN) model aimed at detecting human movement in crowded environments, thereby improving the accuracy of identifying potential risks. However, both methods lack real-time alerting capabilities, which are essential for timely interventions in critical situations. Addressing this gap, the proposed system integrates live-stream processing with automatic suspicious activity detection while simultaneously generating video evidence for verification of detected threats.

Kranthi Kumar *et al.* [1] introduced a machine learning method that identifies human activities in security videos to enhance detection reliability. Sale *et al.* [2] developed a CNN-based anomaly detection framework tailored for municipal environments, contributing to improved public safety monitoring. Despite their achievements, these systems are primarily focused on recorded video data, failing to address delays in threat detection and response, which remain crucial for real-time surveillance.

Recent studies have emphasized the integration of the Internet of Things (IoT) with vision-based modalities to create intelligent security systems. Rehman and his team [7] designed an IoT-based surveillance system that unifies various sensors to strengthen smart city security infrastructures. Chen [6] proposed an AI-driven security solution that employs machine learning algorithms to detect abnormal movements in high-risk zones. While these approaches enhance detection capabilities, they overlook the reduction of response times—a critical requirement for effective real-time security operations. In contrast, our system applies deep learning techniques to analyze video frames continuously, ensuring prompt notification to security personnel upon detection of potential threats.

Advancements in anomaly detection include Maqsood *et al.*'s [10] use of a 3D CNN model for analyzing recorded footage to improve threat identification, and Mittal *et al.*'s [11] development of SAVCHOI, a semantic video captioning system that examines human-object interactions for more interpretable anomaly detection. Bouhlela *et al.* [9] further contributed by designing a UAV-enabled surveillance framework capable of behavioral analysis and facial recognition. Meanwhile, Radhika and Muthukumaravel [12] combined traditional surveillance methods with deep learning to achieve higher accuracy rates. However, the majority of these studies emphasize post-event analysis rather than immediate detection, leaving a gap for systems capable of live-stream anomaly identification and instant alerting.

Our proposed approach builds upon these advancements by employing advanced deep learning models that integrate real-time monitoring, automatic anomaly detection, and immediate alert notifications. This system ensures faster response times, minimizes dependency on manual video reviews, and enhances overall security effectiveness. By applying this solution across crowded public spaces, critical infrastructure, and smart cities, the framework delivers improved situational awareness. Future developments in this area may benefit from multimodal sensor fusion, predictive analytics, and adaptive learning mechanisms to further strengthen real-time anomaly detection and scalability across diverse security applications.

Suspicious behavior detection, and notification alerts. This ensures faster response times, reduces reliance on manual video analysis and increases security effectiveness across the board. Bringing decision-making and situational awareness to crowds, public places, critical infrastructure, and smart cities through integrated live-streaming video surveillance with AI for threat detection. Advanced models that integrate both human-like reasoning and machine intelligence will be crucial for increasing the accuracy of detection while enabling scalability across a breadth of security applications, and for adopting machine learning across current solutions. Future developments might use multimodal sensor integration, predictive analytics, and adaptive learning models.

PROPOSED METHODOLOGY

The system combines advanced analytical methods through which learning algorithms execute event evaluation and dangerous item identification tasks. CNNs process object detection through labeled datasets when detecting anomalies. **Pressing** events linked to robbery occur due to the combination of temporal models with grouping methods in event analysis systems. The system ensures reliable performance by merging algorithmic learning approaches with visual computing backgrounds as shown in **Fig. 1**.

***Corresponding Author:** Jyoti Neeli

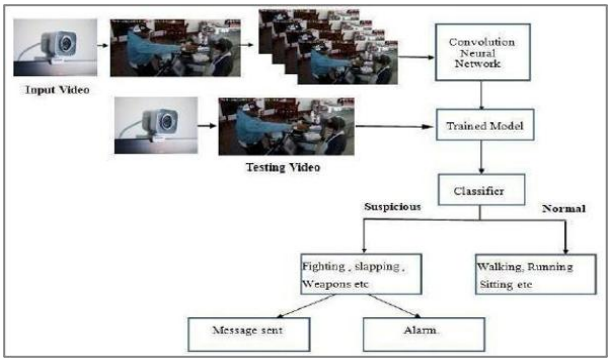

Fig. 1. System architecture

The intelligent surveillance system depends upon machine learning to evaluate video information and spot real-time anomalous behaviors as **Fig. 1** displays. Security equipment begins video recording at different public areas including streets and workplaces through their CCTV cameras. A video examination requires removing video frames to study them individually. Next, the deep learning model Convolutional Neural Network (CNN) functions to process frames received from the first processing step.

Through trained movement patterns and interaction analysis, the CNN determines between standard and deviant behaviors since it analyzed human activities across multiple domains when being trained.

After training, the model serves to detect observed behaviors in fresh video data through its constant evaluation of streaming video frames. The categorization divides behaviors into two sets where suspicious activities include fighting along with slapping or sitting and typical activities include walking and running and using a weapon, armed with a weapon. Automated alerts are triggered to authorities for quick response and an alert system notifies security personnel about any concerning activities. The surveillance system tracks activities without alerting anyone only when no dangerous activities are present. Multiple smart city projects together with automated security systems and crime prevention technology find significant value through this AI-driven surveillance system. The security system represents a modern addition to current defenses since it enhances community security and reduces human-operated monitoring through its immediate alert systems and real-time analysis abilities.

The Motion Detection and Alert System Flowchart in **Fig. 2** outlines an in-depth process for motion detection and alarm activation procedures. The system features a continuously operable camera with two key functionalities which include streaming video and motion detection abilities. Frame recording and temporary storage take place within memory during periods when motion detection activation occurs. The system checks acquired frames through pixel-by-pixel comparison with stored reference frames for identifying changes. The system uses RGB (Red, Green, Blue) values to check all the pixels contained within frames. The surveillance system continues its observation while staying in a sedentary state because no changes have been detected. The system uses verification protocols which activate an alert system if the RGB values change.

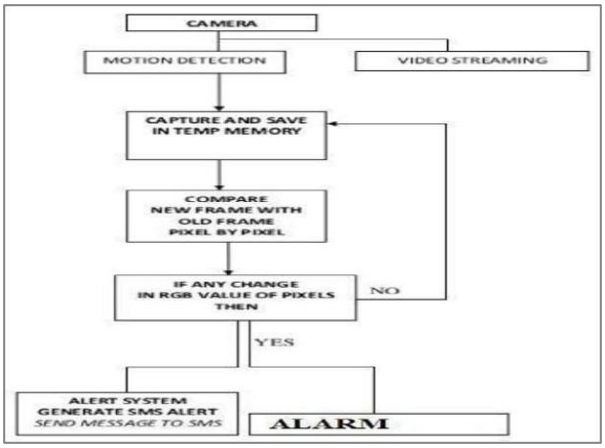

Fig. 2. System Flowchart

This alert system involves sounding an alarm to warn people in the vicinity and sending an SMS notice to the relevant authority. The system ensures real-time security and surveillance by continuing to watch for additional movements after producing the warning. The system is very useful for applications like intrusion detection, security monitoring, and smart surveillance systems because of this automated approach. It guarantees constant observation and prompt action in the event of any movement detected, enhancing safety and security in a variety of settings.

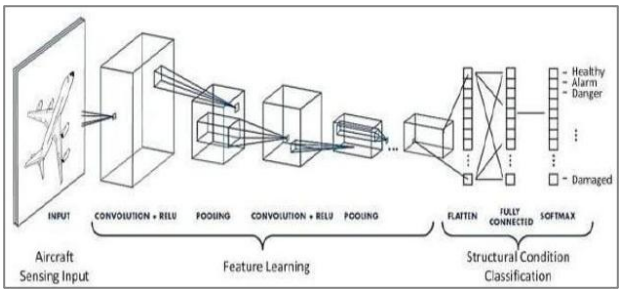

Fig. 3. Neural Network Architecture

Fig. 3 shows a neural network architecture that uses deep learning to classify airplane structural conditions. The first step in the procedure is the aircraft sensing input, which involves taking a picture of an airplane and feeding it into the model. In order to extract important structural elements that aid in detecting potential damage, this input is subsequently sent through several layers of a Convolutional Neural Network (CNN). The input is sent through a number of convolutional layers during the feature learning stage, which apply filters to identify important patterns including edges, textures, and structural elements. A ReLU (Rectified Linear Unit) activation function comes after each convolutional operation, adding non-linearity and allowing the network

***Corresponding Author:** Jyoti Neeli

to understand intricate correlations. Additionally, the network has pooling layers to lower dimensionality while maintaining the highest significant characteristics, increasing computational effectiveness. To improve feature extraction and get the data ready for classification, these procedures are carried out several times.

The model moves on to the structural condition classification phase after the features have been extracted. High-level patterns are discovered by passing the collected data through fully linked layers after it has been flattened into a one-dimensional vector. Based on the identified structural features, the last layer uses a Softmax activation function to divide the aircraft's state into several groups, including damaged, danger, alarm, and healthy. Automated aircraft inspection and maintenance leverage the deep learning approach because this technique provides speed in identifying potential structural damage during inspections. The methodology builds aviation safety and reliability through CNNs for detecting problems early, which reduces the likelihood of structural collapse.

ALGORITHMS USED

YOLOv5

The wide application of YOLOv5 occurs in robotics fields combined with autonomous driving systems while also being used in surveillance operations because of its rapid performance alongside high accuracy and adaptable features.

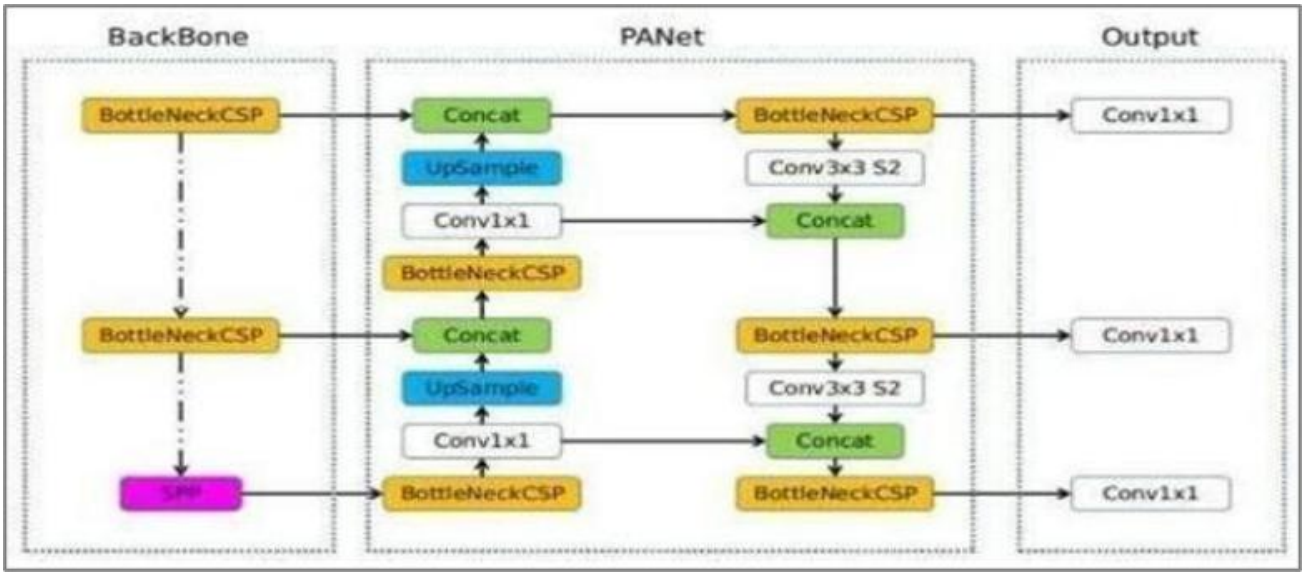

Fig. 4. YOLOv5 Architecture

Vision-based AI jobs requiring fast, precise identification find the perfect solution in YOLOv5 because of its ability to perform effective real-time object detection. YOLOv5, Fig. 4, boosts efficiency through data augmentation while improving dependability through automated anchor change implementations for wide application needs. After feature extraction, PANet applies data processing techniques which focus on feature map combination and enhancement to achieve better detection results. This part employs Conv1×1 layers for channel depth management and Concat layers to blend features from different layers and uses Up Sample layers for higher-resolution detail recovery. By using Bottle Neck CSP layers, the feature representation achieves higher standards, thus leading to more precise object localization and identification. The enhanced features travel through Conv1×1 layers during the last step until detection results appear.

The detection model generates three types of results that include positioning boxes together with object categorization scores and detection confidence scores for reliability assessment. The model applies Non-Maximum Suppression (NMS) to remove redundant bounding boxes as part of its operation, which leads to better accuracy in essential detection results. The YOLOv5 detection flow begins with applying feedforward input images to the feature extraction backbone stage before PANet performs feature fusion operations that result in detection outputs.

YOLOv5 finds broad applications in robotics, autonomous driving, and surveillance control because of its exceptional speed and accuracy capabilities along with its scalability advantages. Vision-based AI jobs requiring fast, precise identification find the perfect solution in YOLOv5 because of its ability to perform effective real-time object detection. YOLOv5 boosts efficiency through data augmentation while improving dependability through automated anchor change implementations for wide application needs.

Backbone Network: YOLOv5 adopts the Convolutional Neural Network (CNN) as its main framework to extract essential characteristics from input pictures. The default

***Corresponding Author:** Jyoti Neeli

backbone structure named CSPDarknet53 represents an optimized version of the Darknet architecture known for its efficient yet precise performance.

Neck: To enhance feature fusion and integration across various scales, YOLOv5 integrates a neck module named PANet (Path Aggregation Network). In order to improve the model's ability to detect objects of varying sizes, this module combines characteristics from different backbone network levels.

The output of the model consists of predictions that include bounding box center coordinates along with width and height measurements coupled with objectness scores and classification probability for separate object types. YOLOv5 maintains a position as a preferred system for workloads that need instant performance alongside precise object recognition because its design focuses on exact identification and rapid feature analysis as well as combining multi-level features.

TESTING

The main purpose of testing is to detect all potential mistakes and weaknesses. The process of testing demands identifying all potential weaknesses as well as vulnerabilities that exist in either systems or products. Tests verify that products along with sub-assemblies and assemblies operate properly with individual parts. The goal of testing is to check that the software system matches user expectations and requirements while functioning dependably along with a low malfunction frequency. The execution of several tests supports complete evaluation to fulfill specific testing criteria.

Types of Testing

Black-Box Testing: Under Black-Box Testing, testers evaluate applications without knowledge of system base functions. The expert evaluates the application during tests without any knowledge about the program code structure or system operational mechanisms. During black-box testing, users examine the application interface with their inputs and output evaluations because they lack knowledge about the system's internal operational processes.

White-Box Testing: White-box testing entails a careful analysis of the code's fundamental structure and logic. For glass testing, commonly known as **open-box** testing, testers require full know-how regarding the internal workings of the application. The evaluation of source code through testing reveals sections with speculatively unreasonable operational results.

DISCUSSION

Video surveillance applications have caused an extreme rise in the need to detect and analyze security and safety events. The development of efficient monitoring systems requires integration between computer vision technology with machine learning capabilities and data analysis practices as well as ethical considerations. The combination of optical flow and background subtraction methods helped but failed when dealing with challenging video monitoring scenarios.

The deployed monitoring methods experienced difficulties detecting abnormal behavior since they failed to handle both moving backgrounds and crowded areas.

Deep learning methodology progress has revolutionized the practices of detecting abnormal activities. The development of strong neural network models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks enables activity recognition to achieve higher accuracy in its results. The strength of deep learning techniques in detecting temporal patterns from video data leads to the identification of normal behavior patterns while simultaneously detecting abnormal movement patterns. Real-time video surveillance obtains considerable benefits from CNNs, RNNs, and LSTM networks because these networks enable quick detection of security-related suspicious activities.

Advanced analytical approaches enable thorough movement observation of human bodies, which helps experts discover faint behavioral symptoms. Video surveillance systems have become better processors of complex changing environments thanks to deep learning methods that lead to increased effectiveness and precise performance. The advancement of surveillance technology through innovative measures produces better, reliable surveillance instruments which improve both threat detection and safety protection in different environments.

RESULTS

Fig. 5. Sign-in Page

The Suspicious Activity Detection System demands users to provide their username along with their password for system entry as shown in **Fig. 5**. System novices have the capacity to register by supplying required personal information including email and password, which creates their account. The system enables users to watch active surveillance feeds and get alerts about suspicious activities once they successfully **log in**.

***Corresponding Author:** Jyoti Neeli

Fig. 6. Unusual event detection in recorded footage

Algorithms with machine learning mechanisms detect unusual behavior in recorded video clips as illustrated in Fig. 6. The system helps security personnel by automatically detecting both interesting and suspicious occurrences, which enhances both security monitoring and surveillance.

Fig. 7. Weapon Detection

The security system detected **a** sword as in **Fig. 7**, **which** would trigger weapon alerts in its monitoring process. The detection is marked by a bounding box as part of the system presentation, which signals possible security threats or unusual activities.

Fig. 8. Weapon detected in live streaming surveillance camera

The security system technology solution demonstrates its ability to detect a pistol as in Fig. 8 during real-time video surveillance of an image. AI-powered threat detection systems generate alerts that include detected threatening images together with warnings which are sent to system operators. The system serves an essential purpose by protecting airport, educational facility, and public-area security while performing quick protective actions. Security staff receive alerts from the detection system at the same time as the system triggers necessary protective systems to prevent potential incidents. Real-time monitoring capabilities together with automatic threat alerts enhance overall safety levels because of this technology implementation.

Fig. 9. Message sent to the owner about the unusual activity

A real-time security system displays the surveillance feed so it can detect weapons and suspicious behaviors in the monitored areas. Through its AI capabilities, the model detects dangerous items such as knives and swords and automatically sends warning alerts to program owners as shown in **Fig. 9**. In addition to weapon detection, the system identifies abnormal events including shoplifting, thus providing the owner with better surveillance management capabilities. The system improves security through fast alert notifications that also deliver identified images, which enables prompt proactive measures to prevent risks.

CONCLUSION

The integration of video surveillance systems into public domains as well as private establishments creates strong security benefits by detecting anomalous activities. Real-time identification and classification of abnormal behaviors become possible through the mutual application of computer vision methods and Convolutional Neural Networks (CNNs) with machine learning algorithms. Security models detect recurring video patterns of potential threats before activating the correct authority notifications system. Security assessment achieved through real-time detection plays an essential role in preventing risks, yet supports performance of security breach responses along with public safety in public events and urban areas and transportation facilities. The operation of machine learning systems depends on significant amounts of well-categorized training data to deliver proper and reliable results when in use. Accurate model performance requires the optimization of all hyperparameters.

The implementation of surveillance systems requires addressing related ethical and privacy issues at hand. The expansion of these technologies demands the implementation of protections between security enhancement and privacy rights. These systems need proper implementation that must include established

***Corresponding Author:** Jyoti Neeli

guidelines and complete transparency and compliance with privacy regulations to function responsibly.

Real-time suspicious activity detection, which uses machine learning, has proven potential to advance public security and safety at its peak. These technologies need proper attention to data quality alongside system deployment along with ethical considerations to achieve effectiveness alongside privacy protection.

FUTURE SCOPE

The predictive functionality of video surveillance systems is expected to experience substantial growth because of emerging modern technology advances in the future. Within the coming years, machine learning models will achieve greater accuracy, which will lead systems to detect various kinds of suspicious behaviors including difficult ones and subtle ones. The reliable detection of suspicious behaviors would increase at the same time false alarms would decrease for mission-critical deployment sites of this technology including airports and public events. Smart technology integration of IoT devices alongside environmental sensors enhances video surveillance abilities because it provides needed data to detect security threats. Edge computing implementation remains vital because it enables processing data inside local networks, thus reducing response delays and boosting reaction times. The collection of better intelligence depends on the combination of video data analysis with traditional monitoring using audio and biometric information. Predictive analytics based on artificial intelligence represents a future security development that allows threat detections before they occur. This approach to safety.

REFERENCES

1. Kumar, K. K., Kumari, B. H., Saikumar, T., Sridhar, U., Srinivas, G., & Reddy, G. S. K. (2022). *Suspicious activity detection from video surveillance*.
2. Sale, M., Patkal, A., Mahale, H., Lavhale, J., & Apsingekar, S. (2022). *Deep learning approach for suspicious activity detection from surveillance video*.
3. Gayathri, M., Meghana, M., Trivedh, M., & Manju, D. (2022). *Suspicious activity detection and tracking through unmanned aerial vehicle using deep learning techniques*.
4. Chen, J. I. Z. (2020). *Smart security system for suspicious activity detection in volatile areas*.
5. Bora, T. S., & Rokade, M. D. (2021). *Human suspicious activity detection system using CNN model for video surveillance*.
6. C. V., Jyotsna, C., & A. J. (2020). *Deep learning approach for suspicious activity detection from surveillance video*.
7. Rehman, T., Saba, M. Z., Khan, R., Damasevicius, R., & Bahaj, S. A. (2021). Internet of Things-based suspicious activity recognition using multimodalities of computer vision for smart city security.
8. Bouhlela, F., Mliki, H., & Hammami, M. (2021). Suspicious person retrieval from UAV sensors based on part-level deep features.
9. Maqsood, R., Bajwa, U. I., Saleem, G., Raza, R. H., & Anwar, M. W. (2021). *Anomaly recognition from surveillance videos using 3D convolutional neural networks* (arXiv Preprint No. arXiv:2101.01073). https://arxiv.org/abs/2101.01073
10. Mittal, A., Ghosal, S., & Bansal, R. (2022). *SAVCHOI: Detecting suspicious activities using dense video captioning with human-object interactions* (arXiv Preprint No. arXiv:2207.11838). https://arxiv.org/abs/2207.11838
11. Radhika, R., & Muthukumaravel, A. (2024). Video surveillance and deep learning: Enhancing security through suspicious activity detection. In *Proceedings of the International Conference on Recent Trends in Computing and Information Technology*.

***Corresponding Author:** Jyoti Neeli