iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** V **Month of publication:** May 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.53392

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue V May 2023- Available at www.ijraset.com*

# Using CNN Algorithm Detection of Suspicious Activity from Video Surveillance

Sushant Shrivastav[1], Sourabh Yadav[2], Devendra Sutar[3], Omkar Dhanwat[4], Prof. Dipali Pawar[5]
[1, 2, 3, 4]*BE Students, Zeal College of Engineering and Research, Pune, Maharashtra, India*
[5]*Assistant Professor, Department Of Computer Engineering, ZCOER, Pune*

***Abstract:*** ***Detecting suspicious human activity involves predicting the positions of body parts or joints from an image or video. This problem has been extensively studied in computer vision for over 15 years. It is crucial because there are numerous applications that can benefit from activity detection. For instance, human pose estimation is used in various areas such as video surveillance, tracking and understanding animal behaviour, detecting sign language, improving human-computer interaction, and capturing motion without markers.***
***Keywords: Real time CCTV footage, Deep Learning, CNN, Detect Human Suspicious Activity.***

## I. INTRODUCTION

Recognizing human behavior in the real world has many practical uses, like smart video surveillance and analyzing shopping habits. Video surveillance is widely used in various places, both indoors and outdoors. It plays a vital role in ensuring security and safety. Security cameras have become a part of our daily lives for protection and peace of mind. In India, e-surveillance is an important focus of the Digital India development program initiated by the government. Video surveillance offers several benefits, such as effective monitoring, cost-effective auditing, and keeping up with new security trends. As crime rates increase, it becomes crucial to identify suspicious activities promptly and take necessary precautions. Many cities and towns have surveillance systems in place that continuously gather data. With the large amount of surveillance data being collected, there is a higher chance of detecting suspicious activities. However, these tasks usually require human supervision because they are complex for artificial intelligence to handle and require significant resources. One approach to simplify automation is by breaking down complex tasks into smaller parts and detecting patterns that may indicate potential crimes. Our focus is on identifying two main potential indicators of criminal behavior using our models.

## II. LITERATURE SURVEY

According to [1], The method described here is based on the idea that when something unusual happens, the recent frames of a video will look noticeably different from the older frames. We use a special model that has two parts: one part focuses on extracting important features from the video frames, and the other part learns the patterns and changes over time. To train this model, we use video data that only contains normal scenes.

The goal is to minimize the difference between the input video and the reconstructed output created by the trained model. Once the model is trained properly, we expect that normal videos will have a small difference between the original and reconstructed versions. On the other hand, videos with abnormal scenes will show a larger difference.By comparing the reconstruction error of each tested input video with a threshold value, our system can detect when something abnormal occurs. Essentially, if the error surpasses the threshold, we can identify it as an abnormal event.

According to [2], Several recent methods suggest the use of CNNs to address specific imaging tasks. These CNN architectures share similarities with the ODP framework but have different design motivations. For example, Schuler et al. introduced a network for deblurring that incorporates a single, predetermined deconvolution step followed by a CNN that is learned. This can be seen as a similar concept to an initial step in the ODP framework but with a different starting point.

According to [3], Sultani et al. employed a comprehensive model for anomaly detection based on the deep Multiple Instance Learning (MIL) framework. They assessed the performance of their method by utilizing frame-based receiver operating characteristics (ROC) curves and calculating the corresponding area under the curve (AUC). The experimental findings demonstrated that the MIL approach for anomaly detection yielded substantial enhancements in detection performance compared to existing state-of-the-art methods.

According to [4], This research paper addresses the challenges associated with real-world images and their impact on object detection. The study focuses on training the YOLO model using degraded images and examines the outcomes. The findings indicate that training the model with degraded images enhances its ability to learn additional features and effectively handle complex environments. This training approach significantly improves the average precision of object detection, enhances generalization capabilities, and increases overall robustness of the model.

According to [5], This research paper suggests the development of a mobile application aimed at detecting abnormal crowd behaviour and managing it effectively. The application is designed to work in conjunction with an IP camera connected to a server-side application. The IP camera is responsible for detecting any suspicious activity, while the server-side application utilizes the Social Force Model (SFM) algorithm to assess the crowd level at the entrance of public or private locations. The gathered information is then transmitted to the user via the mobile application.

According to [6], In the initial studies on weapon detection, the emphasis was on analyzing x-ray and infrared images to identify hidden weapons. One study employed color-based segmentation techniques to differentiate objects and utilized the Harris interest point detector and FREAK descriptor to detect guns in the segmented images. Similarly, various approaches have been developed to detect cyberbullying, including the identification of deceptive phishing in instant messages for text- based messages.

According to [7], There are still some drawbacks in the detection process of suspicious words in instant messaging (IM) and social networking sites (SNS). This is because these platforms often contain shorthand or abbreviated words, making it difficult for regular surveillance tools to identify suspicious messages, allowing the concealment of such content. To address security concerns, an Advanced Motion Detection (AMD) algorithm was utilized to identify unauthorized entries in restricted areas. The algorithm employed a two-phase approach. In the first phase, objects were detected using background subtraction, and then the objects were extracted from the sequence of frames. The second phase involved the detection of suspicious activities. The key advantages of this system were its real-time video processing capabilities and low computational complexity. However, the system had limitations in terms of storage capacity, and it could be further enhanced by implementing high-tech video capture methods in the surveillance areas.

According to [8], This exploration paper introduces the discovery and recognition of suspicious mortal exertion in surveillance vids, which is an on going field of study in image processing and computer vision. Visual surveillance plays a pivotal part in covering mortal conditioning in colourful sensitive and public locales, including machine stations, road stations, airfields, banks, shopping promenades, seminaries, sodalities, parking lots, and roads. The end is to help and address colorful issues similar as terrorism, theft, accidents, illegal parking, vandalization, fights, chain swiping, crime, and other potentially suspicious conditioning.

## III. REQUIREMENT SPECIFICATION

### A. *Hardware Requirement*

1) *Processor (CPU):* Intel 5 or above processor is needed. A multi-core processor or a dedicated high-performance processor is recommended. The CNN algorithm can be computationally intensive, especially when dealing with large datasets or complex feature spaces.
2) *Memory (RAM):* The amount of RAM required depends on the size of your dataset and the complexity of the CNN model. Though a minimum of 4GB of memory is required.
3) *Storage:* A minimum of 40GB of hard disk is required. Sufficient storage is required to store the dataset, feature vectors, and the CNN model.

### B. *Software Requirement*

1) *Programming Language:* You'll need a programming language that supports CNN implementation and offers libraries or packages for machine learning. We have used Python programming language.
2) Integrated Development Environment (IDE): An IDE can provide a user-friendly development environment with features like code editing, debugging, and project management. We have used PyCharm, Anaconda (with Spyder).

## IV. METHODLOGY

Identifying suspicious human behaviour has received considerable attention in the field of computer vision for more than 15 years. Our aim is to leverage neural networks to address these challenges. The recognition of suspicious activity in surveillance videos is a dynamic research field within image processing and computer vision. Our methodology involves the continuous monitoring and analysis of CCTV footage in real-time.

By analyzing the footage, we can generate actionable commands to alert the appropriate authorities when potential incidents are detected. The model takes video inputs and produces outputs that indicate the presence of suspicious activity. In such instances, an alert message is promptly displayed to administrators or security analysts.

### A. *Convolutional Neural Network (CNN) Algorithm*

A Convolutional Neural Network (CNN) is a deep learning algorithm widely used for analyzing images and videos. CNNs are specifically designed to automatically learn and extract patterns and features from raw input data, such as images.

In CNNs, convolutional layers are utilized, where a set of filters (also known as kernels) are applied to the input image by sliding over it. Each filter performs a dot product between its weights and a small local region of the input image, generating afeature map. Multiple filters are employed to capture different features present in the image. Pooling layers are used to reduce computational complexity and enhance the network's resilience to variations in the input. These layers aggregate the information from the feature maps, reducing their size. Fully connected layers come after the convolutional and pooling layers. The final feature maps are flattened into a one- dimensional vector and fed into one or more fully connected layers. These layers compute a weighted sum of the inputs, followed by an activation function, to produce the desired output.

### B. *Python*

Python is a flexible programming language that can be highly valuable in detecting potentially suspicious behavior in surveillance videos. Python offers a variety of libraries and frameworks that make video processing more accessible. These libraries provide extensive features for reading, manipulating, and analyzing video frames. Python also offers a vast collectionof tools, libraries, and frameworks that empower developers to implement intricate algorithms and models for detecting suspicious activity in surveillance videos. Its user-friendly nature, extensive community support, and diverse ecosystem make it a favoured option for building surveillance systems with advanced video analysis capabilities.

### C. *Spyder*

Spyder is an integrated development environment (IDE) bundled with the Anaconda distribution, specifically designed for scientific computing and data analysis in Python. Its primary function within Anaconda is to provide a user-friendly environment for writing, running, and debugging Python code, with a particular focus on tasks related to data science and numerical computing.

Spyder offers a comprehensive code editor with a range of useful features such as syntax highlighting, code completion, and code introspection. It also includes an interactive Python console where code can be executed interactively, allowing for quick experimentation and exploration. One of Spyder's key strengths is its seamless integration with the Anaconda ecosystem. It harmoniously works with the extensive set of pre-installed data science libraries and tools provided by Anaconda, enabling users to leverage the power of these libraries directly within the IDE. This integration enhances productivity and facilitates efficient data analysis workflows.

### D. *SQLite database*

1) The main purpose of SQLite database is to offer a lightweight, self-contained, and serverless database solution thatcan be seamlessly integrated into applications. It enables structured data storage and management in a relational database format. Users can create tables, define their structure (columns and data types), and manipulate the data stored within them.
2) SQLite adheres to the ACID principles (Atomicity, Consistency, Isolation, Durability), ensuring reliable and consistent data transactions. It supports transactions, enabling multiple database operations to be grouped together, guaranteeing data integrity and consistency.
3) SQLite finds extensive usage across various applications, including mobile apps, desktop software, web browsers, IoTdevices, and embedded systems. Its simplicity, portability, and efficiency make it a popular choice in scenarios wherea lightweight, self-contained, and user-friendly database solution is required.

## V. IMPLEMENTATION

1) *Dataset Preparation:* Collect a diverse range of video data containing instances of both suspicious and non-suspicious activities. Annotate the collected data by marking the timestamps or frames where suspicious activities occur. Assign appropriate class labels or categories to differentiate between suspicious and non-suspicious instances. This annotation process can be done manually or through crowd-sourcing platforms, depending on the dataset's size and complexity

2) *Feature Extraction:* Extract motion-related features like optical flow, speed, direction, or changes in position to identify abnormal movements or sudden activity changes. Analyze object shape and contour for valuable information, such as object size, aspect ratio, elongation, or convexity.
3) *Data Pre-Processing:* Pre-process the extracted features by extracting individual frames from video segments. Frames serve as input for analysis. Perform cleaning operations to address noise, artifacts, or inconsistencies in the frames, utilizing techniques like denoising, image enhancement, or filtering to improve frame quality.
4) *Model Training:* Split the pre-processed data into training and validation sets. Train the selected model using labeled data and extracted features.
5) *Model Evaluation:* Evaluate the trained model on an independent test set to measure its performance. Common evaluation metrics include accuracy, precision, recall, F1-score, and area under the ROC curve (AUC-ROC).
6) *Deployment and Monitoring:* Design a scalable and available system, especially if real-time video feeds from multiple surveillance cameras need to be processed. Consider implementing load balancing, containerization, or distributed computing frameworks to ensure scalability and fault tolerance. Implement proper security measures to protect the system and sensitive data. Continuously monitor the performance of the machine learning models used for suspicious activity detection, tracking metrics like accuracy, precision, recall, and false positives/negatives. Monitor model drift over time and retrain or update models as needed.

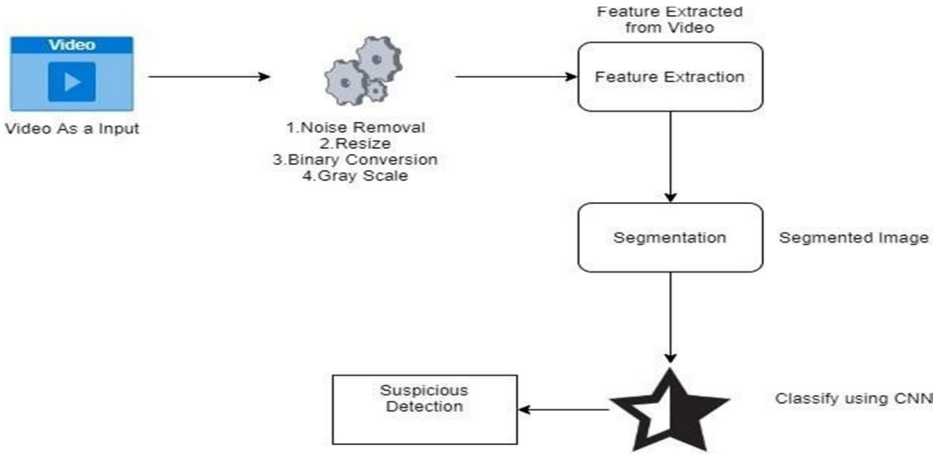


Fig 1: System Architecture

## VI. RESULTS AND DISCUSSIONS

We have to add the images .and it will give us in result that of attack is detected.

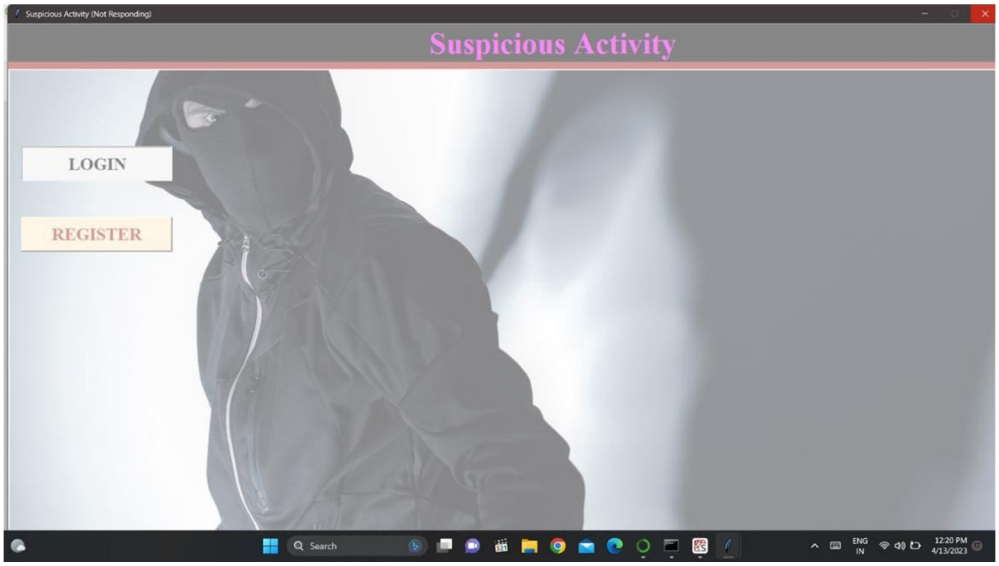


Fig 2: Login Page

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue V May 2023- Available at www.ijraset.com*

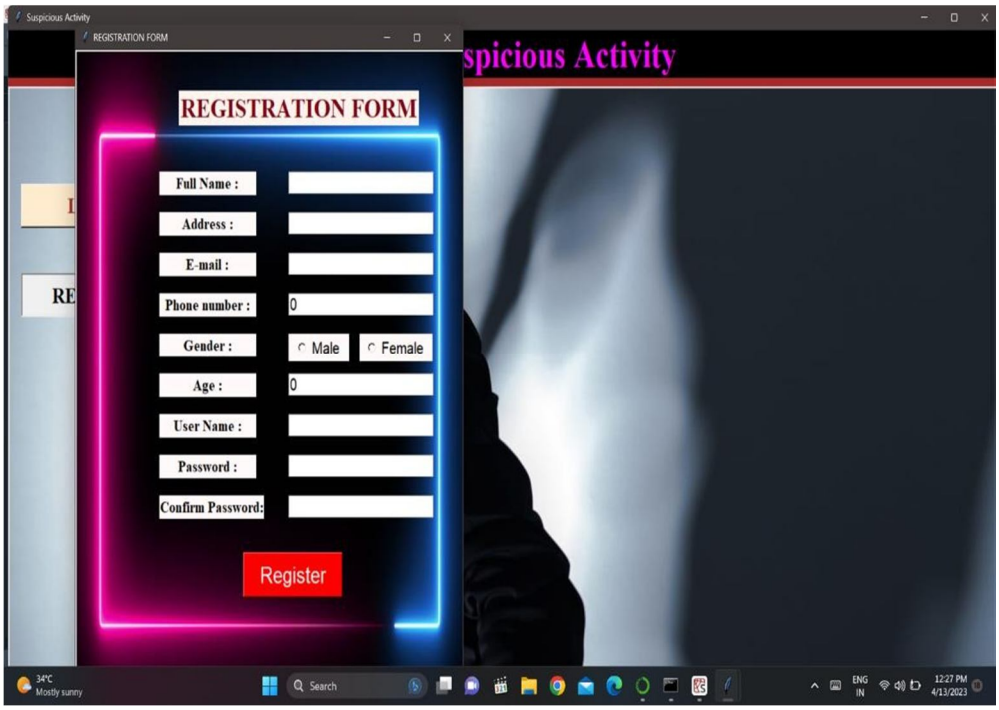


Fig 3: Interface of Registration Form

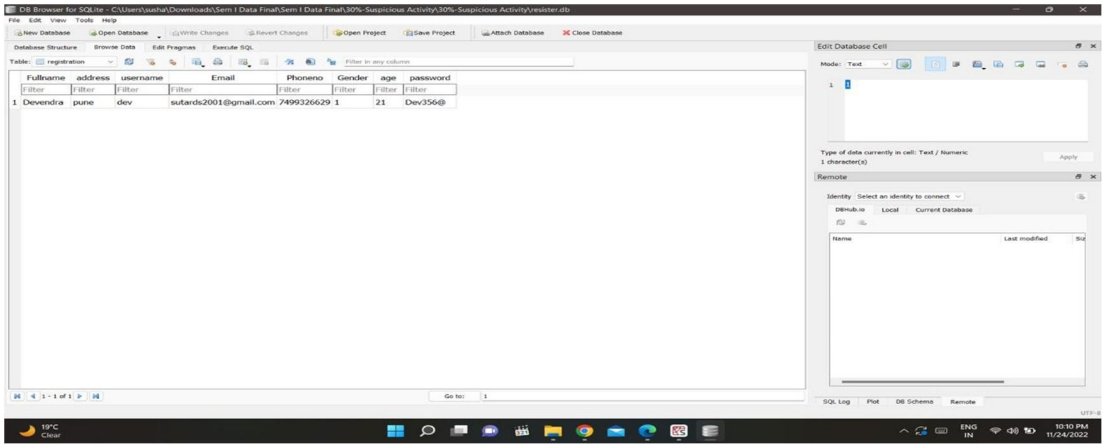


Fig 4: SQlite Databse

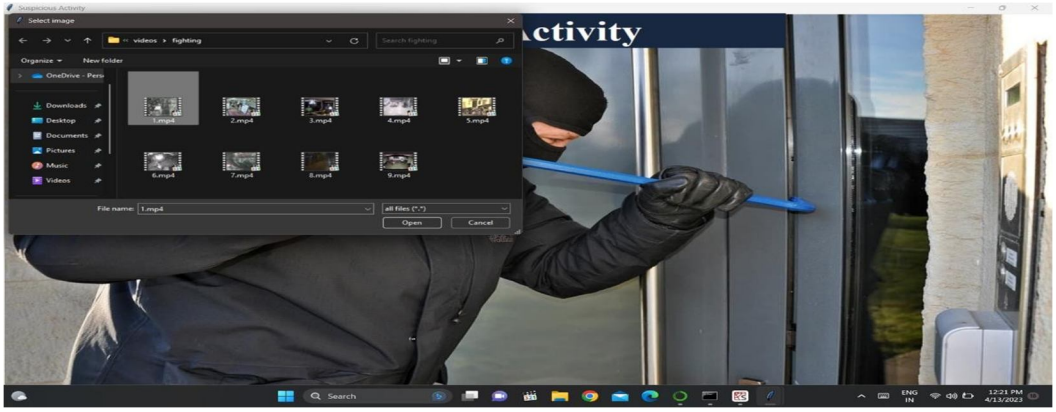


Fig 5: Take Input Video

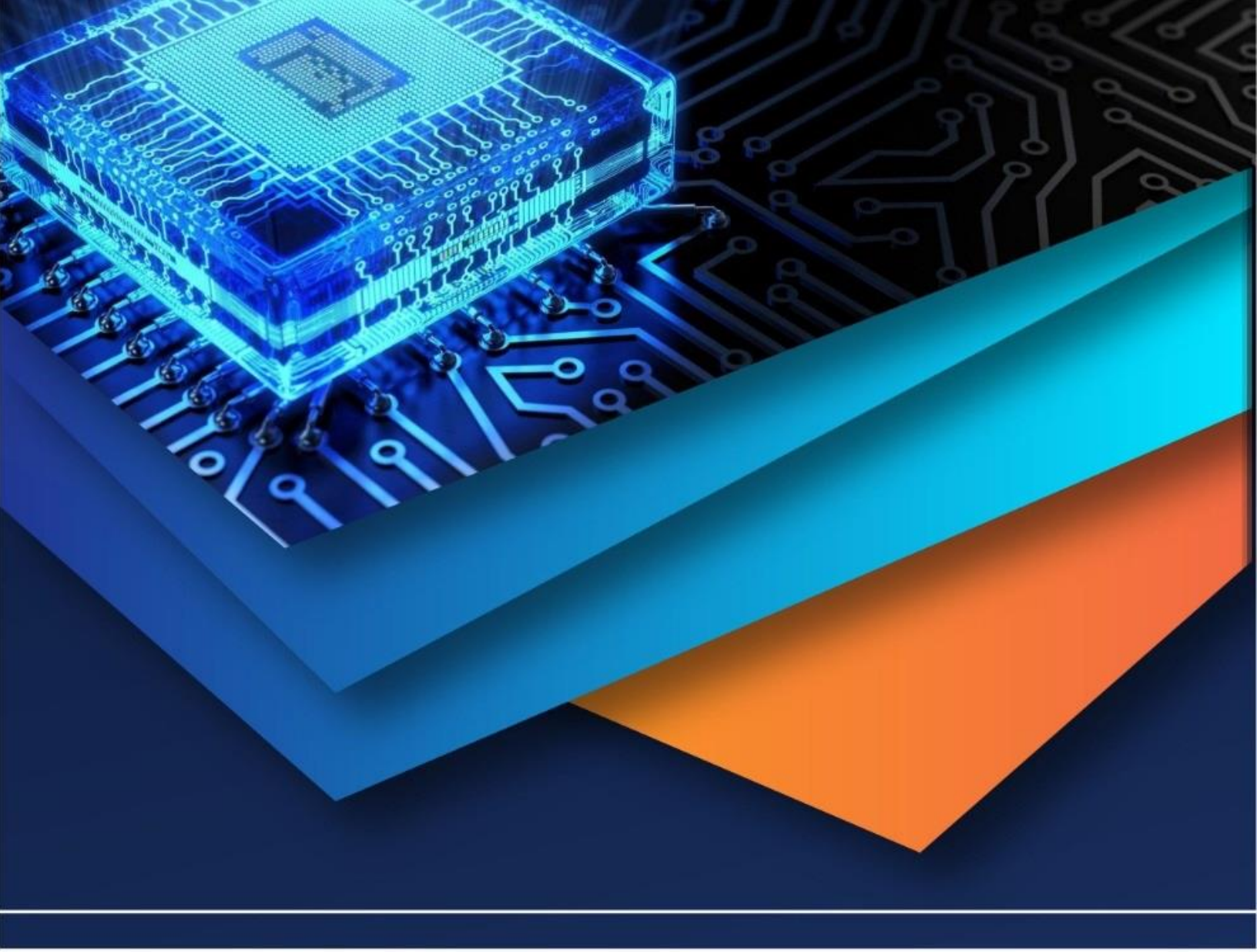

Fig 6: Suspicious Activity Detection

## VII. CONCLUSION

Implementing a real-time CCTV footage processing system for detecting suspicious activity can significantly improve security and reduce the need for human intervention. Notable advancements have been achieved in the domain of identifying suspicious human behavior, allowing for broader applications of this technology. Additionally, ongoing research in related areas like activity tracking can further enhance its practical implementation in various fields.

## REFERENCES

[1] Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A.K., Davis, L.S.: Learning temporal regularity in video sequences. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 733–742 (June 2016)
[2] C. Schuler, H. Burger, S. Harmeling, and B. Scholkopf. A machine learning approach for non-blind image deconvolution.
[3] In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1067–1074, 2013
[4] W. Sultani, C. Chen and M. Shah, ”Real-World Anomaly Detection in Surveillance Videos,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 6479- 6488
[5] Kamthe, U. M., Patil, C. G. (2018) ”Suspicious Activity Recognition in Video Surveillance System” 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).
[6] Liu, C., Tao, Y., Liang, J., Li, K., Chen, Y. (2018) ”Object Detection Based on YOLO Network” 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)
[7] Rohit Kumar Tiwari and Gyanendra K. Verma, “A Computer Vision based Framework for Visual Gun Detection using Harris Interest Point Detector”, Procedia Computer Science, vol 54, p. 703 - 712, 2015
[8] P.Bhagya Divya, S.Shalini, R.Deepa, Baddeli Sravya Reddy,“Inspection of suspicious human activity in the crowdsourced areas captured in surveillance cameras”,International Research Journal of Engineering and Technology (IRJET), December 2017.
[9] Prof. Dipali Pawar, Omkar Dhanwat, Sushant Shrivastav, Devendra Sutar, Sourabh Yadav
[10] “ Suspicious Activity Detection from Video Surveillance Using CNN”,International Journal of Research in Applied Science and Engineering Technology (IJRASET), May 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY