International Journal On Engineering Technology and Sciences – IJETS
ISSN (P):2349-3968, ISSN (O): 2349-3976 Volume XI - Issue III, March – 2024

Suspicious Crowd Activity Detection And Localization Using Computer Vision And CNN

[1]Prof. Mohazzeba Tanveer Raza, [2]Rajesh L, [3]Nandish V, [4]Manoj B M, [5]Rajith A N
[1]Assistant Professor, [2,3,4,5] UG Student, Department of Computer Science and Engineering,
HKBK College Of Engineering, Bengaluru , India.

Abstract—This paper introduces a significant applications of human suspicious activity recognition that is termed as anomaly detection. Addressing the pressing concern of individual safety in society. The alarming rate of criminal activity at banks, airports, temples, parks, sports venues(stadium), hospitals, and shopping centers has created a strong need for smart vision-based surveillance systems. These systems can be used for a variety of human activity recognition applications, such as patient fall detection, irregular pattern recognition, or human- computer interaction. . In public spaces, suspicious behavior can be dangerous and lead to significant casualties. There are a number of systems that have been developed where motion or pedestrian detection occurs based on video frame acquisition, but those systems lack the intelligence to recognize suspicious activity even in real time. Real-time identification of scammer situations from video surveillance is necessary for prompt and effective management to prevent any casualties. The proposed system aims to develop a technique that can automatically detect suspicious activity using computer vision, with a focus on identifying suspicious activities. The framework of Convolutional Neural Network is utilized for processing of images and videos. This system presents information at a level of pixels to make it simple to understand and recognize the real situation.

Index Terms—Convolutional Neural Network , Computer Vision, Suspicious Activity , Video Surveillance.

I. INTRODUCTION

CCTV surveillance is the most important and effective security feature that a building can have in the modern world. The most common method of stopping and identifying undesirable activity is to install CCTV in places like hospitals, universities, shopping centers, etc. Recognizing human activity is helpful in numerical contexts, such as identifying unusual activity in security systems. Surveillance cameras for video analysis are widely used as a result of the rising demand for security. CCTV cameras are installed in many businesses to monitor employees and their activities. One of the most crucial aspects of surveillance video analysis is identifying anomalous behavior[1]. Examining human behavior and phys- ical activity more closely is made possible by the classification of criminal activity in humans. It is possible to determine whether someone is engaging in legal or illegal activities based on their behavior, attire, and weapons. Humans can interact with society through regular or safe activities. emotions A person's basic daily activities include reading books, using a phone, eating, drinking, and walking. converse with one another, write, and sleep. It's simple to find out about activities like eating, writing, and sleeping. Human Suspicious Activity Recognition via Real-Time Video and Transfer Learning for Deep Learning In contrast, other actions like fighting, theft, attacking with a gun or knife, hiding a suspicious person, and chain snatching are really difficult[2].The proposed system aims to use CCTV camera clips to monitor human behavior on campus and gently alert authorities to any suspicious incidents. The main components of intelligent video monitoring are event detection and human behavior reco The field of image processing and computer vision is actively researching the detection of suspicious human activity and fight detection from using video as an input to a system. Visual monitoring allows for the observation of human activity and behavior in public and sensitive spaces like bus stops, shopping centers, banks, airports, train stations, and airports. to stop criminal activity, terrorism, theft, mischief, illegal parking, fighting, chain snatching, and other dubious activities like fight recognition. Since it can be challenging to constantly monitor public spaces, intelligent video surveillance that can identify normal and abnormal activity, generate alerts, and send messages to administrators is required.

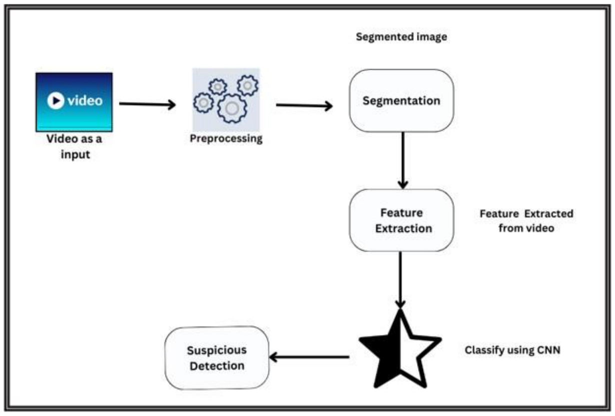


Fig. 1. Surveillance process

[4]. Employing surveillance cameras necessitates the integration of computer vision technologies to analyze extensive amounts of video data. An important application in this domain is the detection of unusual activities within captured scenes. The challenging task of localizing and detecting anomalies in video analysis is addressed in this paper. Our work proposes and assesses a novel approach for detecting suspicious activities. We introduce and investigate a modified pretrained convolutional neural network (CNN) designed for the detection and localization of anomalies within video frames[5]. Image processing and computer vision researchers are continuously concentrating on the identification of questionable human activity from surveillance footage. When it comes to keeping an eye on how people behave in public and sensitive spaces like bus stops, train stations, airports, banks, retail centers, parking lots, schools, and colleges, visual surveillance is essential. The goal is to stop illegal activities such as fighting, chain snatching, vandalism, theft, accidents, and terrorism. Continuous monitoring of public places is challenging, necessitating intelligent video surveillance capable of categorizing human activities as either typical or unusual and generating alerts accordingly[6]. Recently, the installation of CCTV cameras in various public locations has become preva- lent, serving the purpose of overseeing potential suspicious activities. The vigilant monitoring of human behavior in com- munal areas is of utmost importance to thwart terrorism, theft, robbery, accidents, riots, chain snatching, crime, and other suspicious activities[7]. Our initiative involves developing an application to identify anomalous activities among people in public spaces in real-time. This application is specifically tailored for surveillance in areas like malls, airports, and railway stations, where there is a potential risk of incidents such as robbery or shooting attacks. To achieve this, we will harness the power of deep learning and neural networks to train our system. The resulting model will be integrated into both mobile and desktop applications, capable of processing real-time CCTV footage as input. In the event of detecting suspicious poses, the application will promptly alert the administrator, ensuring a swift response to potential security threats. The focus of our approach lies in the identification of human body parts and the possible tracking of their movements to pinpoint anomalies in human activity[8].

RELATED WORK

[1] The research work Titled ”Surveillance-based Suspicious Activity Detection: Techniques, Application and Challenges” proposed system for makes use of a hierarchical method based on object motion characteristics to identify suspicious activity. The Semantic approach categorizes suspicious activity, motion features and temporal data classify events, background subtraction detects objects, and correlation technique tracks them. The semantic approach increases efficiency and decreases computational complexity.[2] The research work Titled ”Real-Time Video based Human Suspicious Activity Recognition with Transfer Learning for Deep Learning” Proposed system for A TL-HAR framework based on transfer learning strategies was presented in Preprocessing, recognition, and pre-training are its three primary stages. Three models are pre-trained to modify network weights using a generic dataset. Using a realistic dataset, this pre-trained network is used to identify human activities.

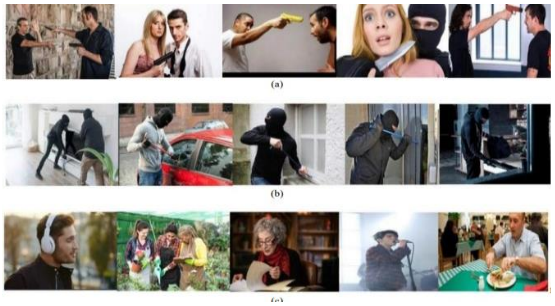
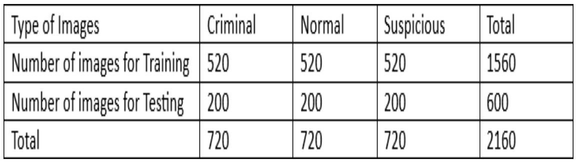


Fig. 2. Datasets: (a) criminal activity (b) suspicious activity (c) normal activity

1. Kaggle Dataset



Type of Images	Criminal	Normal	Suspicious	Total
Number of images for Training	520	520	520	1560
Number of images for Testing	200	200	200	600
Total	720	720	720	2160

Fig. 3. Dataset

[3] The research work Titled ”Suspicious Activity Detection from Surveillance Video using Deep Learning” Proposed system for One method for identifying odd activities in video footage is to track individuals. In order to do this, a background subtraction technique is first used to identify human presence in the video. CNN then extracts the features and feeds them into DDBN (Discriminatory Deep Belief Network). Labeled videos of suspicious events are also fed to DDBN, from which features are also extracted. Features taken from CNN-labeled sample videos and DDBN-classified suspicious actions are compared. This technique allows for the detection of a variety of suspicious activity in a given video.[4]

The research work Titled ”Suspicious Human Activity and Fight Detection using Deep Learning” Proposed system for so many applications could profit from suspicious activity, this is crucial. Applications such as video surveillance, advanced human-computer interaction, sign language detection, human motion tracking and behavior comprehension, and marker low motion capturing, for instance, use human suspicious activity. Low-cost depth sensors have drawbacks, such as being restricted to indoor use, and it is challenging to estimate human pose from depth images due to their noisy depth information and low resolution. Consequently, we employ neural networks to prevent these issues.[5] The research work Titled ”Suspicious Activity Detection through CCTV” This project is a surveillance system that uses video to identify, follow, and keep an eye on suspicious behavior. This paper’s advantage is its ability to monitor multiple screens at once without experiencing the drawback of losing focus. increases the efficacy and efficiency of operations. Still, its inability to distinguish between objects with similar looks is a limitation.[6] The research work Titled ”Human Suspicious Activity Detection using Deep Learning” proposed system for it has to do with recognizing the various body parts of people and maybe following their movements. AR/VR, gesture recognition, game consoles, and other devices use it. In the beginning, human movement was detected by inexpensive depth sensors, also known as motion sensors, in game consoles. These sensors can only be used indoors, though, and it is challenging to infer human activity from depth photos due to their poor resolution and noisy depth data. They are therefore not a good choice for detecting suspicious activity. [7] The research work Titled ”Real Time Video Based Human Suspicious Activity Recognition Using Deep Learning” Proposed system for Numerous investigations have previously been conducted utilizing Human Suspicious Activity Recognition methodologies. Due to the intricate human activity or behavior patterns and identification with machine learning techniques, it is still a difficult task. Recognizing Human Suspicious Activity involves four primary steps, which are as follows: (a) locating the activity region; (b) obtaining the intrinsic values (features) from the area; and (c) going through the normalization process to get rid of the dataset’s significant value fluctuations. (d) Methods of classification for identifying actions or behaviors.[8] The research work Titled ”Suspicious Activity Detection Network For Video Surveillance Using Machine Learning. ”Proposed system for the definition of anomalous events in lengthy video sequences can be ambiguous, making it difficult to automatically identify them. Our method for solving the issue is to train generative models with limited supervision that can recognize abnormalities in videos. Our suggestion is to utilize end-to-end trainable composite Convolutional Long Short-Term Memory (Conv-LSTM) networks, which can forecast a video sequence’s progression using a limited number of input frames. Lower regularity scores are obtained as the predictions with abnormal video sequences diverge from the original sequence over time. Regularity scores are obtained from the reconstruction errors of a set of predictions. The models make use of a composite structure and investigate how ”conditioning” affects the acquisition of more insightful representations. Automating the identification of unusual occurrences in lengthy video clips Based on the reconstruction and prediction accuracies, the optimal model is selected.[9] The research work Titled ”Suspicious Activity Detection Using Convolution Neural Network” Proposed system for Long Short-Term Memory Networks with Predictive Convolution.

Because it's unclear exactly what constitutes an anomalous action in a lengthy video series, they proposed an automated method for detecting it. In order to solve the issue, the authors trained generative models that use restricted supervision to identify anomalies in videos. complex Convolutional Long Short-Term Memory (Conv-LSTM) networks that can be trained end-to-end and projected to predict how a video sequence will unfold from a small number of input frames. [10] The research work Titled ”Suspicious Actions Detection System Using Enhanced CNN and Surveillance Video” Proposed system for This study's operational mechanism involved creating three-dimensional (3D) object-level data for tracking individuals and luggage in real-time public transportation environments. A wide range of behaviors related to the security of public transportation were trained. Fighting, lingering, passing out, and stealing items from open datasets were among the behaviors. The exceptional performance and minimal computational complexity led to the real-time blob matching technique being experimented with. Its limitation was that it was limited to surveillance in public areas. Regarding the peculiar movement of human tracking and detection.[11] The research work Titled ”AI Suspicious Activity Detection using Human Pose Estimation” Proposed system for This work's objective was to find noticeable or unusual occurrences while conducting video surveillance. The advance motion detection (AMD) method was used to find an unauthorized entry into a restricted area. Objects were extracted from the frame sequence and recognized using background subtraction in the first stage. The detection of suspicious activity is the second stage. This system's low computational complexity and real-time video processing capabilities were its main advantages. On the other hand, the system has limited storage capabilities and does not support high-tech video recording modes in surveillance areas. [12] The research work Titled ”Suspicious Activity Detection in Surveillance Video using Discriminative Deep Belief Network” Proposed system for tracking and identifying several moving objects with a single camera. Red-green-blue (RGB) color background modeling is used in the suggested method to extract moving regions. Moving objects are grouped using blob labeling. This method's quick computation times and resistance to environmental influences make it ideal for use in real-time surveillance systems. RGB BM with a new sensitivity parameter was used to extract moving regions, morphology schemes were used to filter out noise, and blob-labeling was used to group the moving objects in order to detect the moving objects.[13] The research work Titled ”Suspicious human activity detection” Proposed system for the people themselves kept an eye on the questionable activity. This was accomplished by designating individuals to constantly watch over people's actions in crowded public areas and look for any unusual activity. Because it is extremely difficult for humans to pay attention for longer than about twenty minutes, this method was a failure. As a result, humans began to be replaced by machines and algorithms in the task of keeping an eye on crowded areas. These days, video analytics is a field that is widely used to identify human activity. [14] The research work Titled ”Toward trustworthy human suspicious activity detection from surveillance videos using deep learning” Proposed system for human action recognition in video interpretation and analysis has become challenging. Monitoring patients in real time is one use of human activity recognition; patients are followed among a group of healthy individuals and subsequently identified based on their unusual behavior. To generate a multi-class abnormal activity detection in individuals and groups, they use video sequences. The foundational CNN model in this study is the You Look Only Once (YOLO) network.[15] The research work Titled ”Suspicious and Anomaly Detection” Proposed system possible to deploy a surveillance system that makes use of Human Activity Recognition techniques to determine whether the target individual is a suspicious person or an average person. It is also anticipated that installing detection systems will serve as a deterrent to criminal activity.

I. PROPOSED WORK

A. Method for Suspicious Activity Detection

It is possible to use convolutional and recurrent neural networks to detect suspicious behavior[16]. The captured frames are categorized into three classes by the trained image processing model, that include:

1. Suspicious Activity
2. Criminal activity
3. Safe or General Activity

B. Yolo algorithm

A system identified as Suspicious Activity Recognition use computer vision techniques to identify and detect potentially suspicious activities in real-time using YOLO (You Only LookOnce) object detection.

The proposed system block diagram is displayed in Figure

1. Typically, the system is made up of the following parts:

1. Video Input: The system receives video streams from security cameras and other sources as input. These video streams have the option of being recorded in advance or inreal time.

2. Pre-processing: To improve their quality and get them ready for object detection, the video frames are preprocessed. Noise reduction, normalization, and resizing are examples of common preprocessing operations.

3. YOLO Model Training: Using the annotated dataset, train a YOLO (You Only Look Once) object detection model. Yolo is renowned for its ability to detect objects in real time.

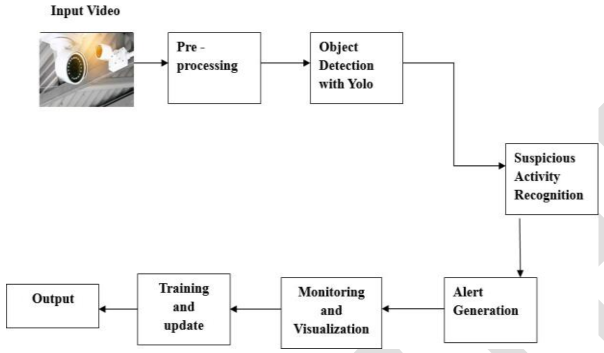

Input Video
Pre - processing
Object Detection with Yolo
Suspicious Activity Recognition
Alert Generation
Monitoring and Visualization
Training and update
Output

Fig. 4. Yolo Architecture

- Create a training environment using the necessary dependencies, such as Darknet (the original framework used by Yolo).

- Get the dataset ready in the format needed for YoLO training, which is usually the labeled bounding box format (also known as the YoLO darknet format).

- Implementing YOLOv4 with CSPDarknet-53, tailored for 80 COCO classes, as per my specifications..

- Using the annotated dataset, begin the training process and modify hyperparameters such as learning rate, batch size, and number of training epochs.

- Track the training procedure, assess the model's effectiveness on a validation set, and save the best-performing model

4. Model Testing: To detect suspicious activity, you can test YOLO model on fresh photos or videos after it has been trained.

- Resize or normalize the test images or frames from the videos to fit the trained YOLO model's input size aspart of the preprocessing step.

- Run the preprocessed pictures or frames through the YOLO model to get class labels and bounding box predictions for objects that are detected.

- Utilize post-processing methods to eliminate false positives and enhance the detection outcomes as required.

- Identify and identify suspicious activities based on the labeled classes using the detection results.

- **Thresholding and Alert Generation:** Based on the confidence scores or other metrics supplied by the YOLO model, establish appropriate thresholds for identifying suspicious activity. It is possible to set thresholds for specific classesor groups of classes.

- Check the detection outputs, like the class labels and bounding boxes, to pinpoint particular object combinations or patterns that suggest questionable activity.

- The notification (or alert) can be delivered through various channels like email, SMS, or system dashboards.

5. Iterative Refinement: Assess your system's performance, taking into account false positive/negative rates and detection accuracy. To enhance overall performance, iterate through the steps of training and testing the model, modifying hyperparameters, gathering additional data, or streamlining the annotation procedure.

C. Data Flow Diagram

Dataset Description: To identify suspicious activity, the LRCN model is suggested. The dataset that this model uses comprises three different kinds of activities.

We have preprocessed this dataset and fed it into our LSTM model. There are three different types of datasets: Fighting, Walking, and Running. They are all made up of one hundred videos. While some of the videos are from YouTube and other sources, some are from Kaggle. Preprocessing of Data :I. Read Video and Label: The videos are read from their corresponding Class folders using an OpenCV library, and the class label is kept inside a numpy array.

ii. Dividing into multiple frames to create a unified sequence: The OpenCV Library is used in this process to read each video, and The videos are divided into smaller segments, each consisting of 30 frames, to facilitate easier analysis.

iii. Resizing: Resizing an image allows you to change how many pixels it has overall. Thus, the frames are scaled down to a 64 pixels for width and 64 pixels for height, so that their sizes match.

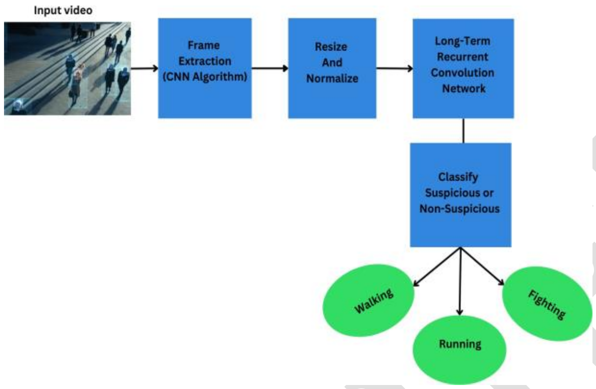

Input video
Frame Extraction (CNN Algorithm)
Resize And Normalize
Long-Term Recurrent Convolution Network
Classify Suspicious or Non-Suspicious
Walking
Running
Fighting

Fig. 5. Data Flow Diagram for Detecting Suspicious Activity

iv. Normalization: By normalizing the images and adjusting their values so that they all fall between 0 and 1, the images are made easier to analyze.

v. Store in Numpy Arrays: A numpy array containing the sequence of thirty resized and normalized frames is provided. as the Model's Input. Split Data for Train Test: Of the data, 75 percentage are used for training. About 25 percentage of the information is used for testing. Model Creation: A deep learning (DL) network, a Long-term recurrent convolutional network (LRCN) is used to proposed the system for suspicious activity detection from video surveillance. A CNN was used by LRCN to process the variable length as visual input. The CNN's outputs were then fed into the stack of recurrent sequence models (LSTMs), resulting in a prediction of the variable length. Over time, the weights of CNN and LSTM are shared, resulting in a representation that can scale to infinitely long sequences. The idea behind LRCN is to analyze videos using two different kinds of computer programs. One sort attempts to decipher what is happening by focusing on the individual images in the video. The other kind looks at how the images change over time in an effort to understand what's happening in the bigger picture. When combined, these two categories of software can assist in comprehending a video and even forecast potential outcomes. LSTM Networks are ideal for processing, categorizing, and forecasting time series data because there can be lags in a time series and between significant events of unknown length. LSTMs were created to address the When training traditional RNNs, one may run into the vanishing gradient problem.

D. System Architecture

Webcam-based Video Surveillance System with Suspicious Activity Detection The architecture depicted in the flowchart represents a video surveillance system that leverages a webcam to capture video footage and a deep learning model to detect suspicious activity. Here's a breakdown of the system's functionalities:

1. Login/Register: The system likely incorporates a login or register functionality, which restricts unauthorized access and enables user management. This step might not be required if the system is designed for local, standalone operation.

2. Choose Feed Type: This section allows the user to choose the type of video feed they wish to view or analyze. Options might include live feed from the webcam or previously recorded videos stored on the disk.

3. Input Video: The system retrieves video input from a webcam or a stored video file on the disk.

4. Preprocessing: The captured video stream undergoes preprocessing to prepare it for analysis by the deep learning model. This typically involves segmenting the video stream into individual frames, potentially resizing the frames for model compatibility, and possibly converting them to a specific color format.

5. Resnet50 Model: A Resnet50 model, a convolutional neural network (CNN) architecture, is employed for real-time frame analysis. Resnet50 is a pre-trained model capable of identifying patterns and features within the frames. It’s likely fine-tuned in this specific application to detect anomalies or suspicious activity within the surveillance area.

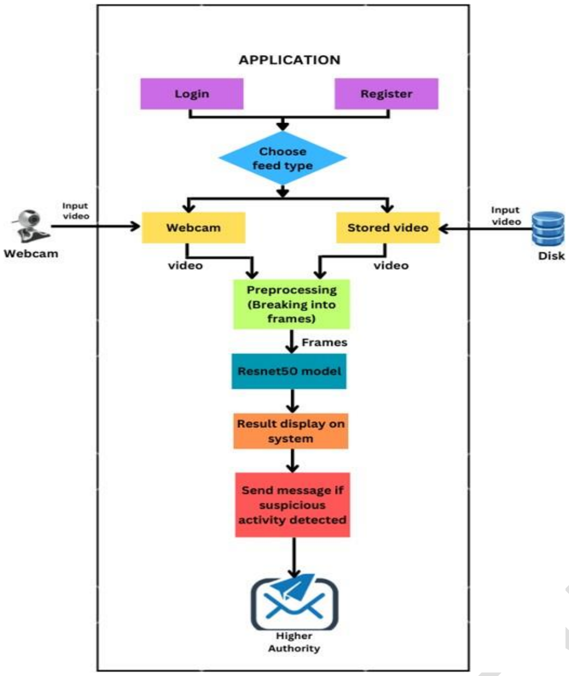


Fig. 6. System Architecture

6. Result Display: The processed video, potentially overlaid with bounding boxes or other visual cues highlighting suspicious activity detected by the model, is displayed on a designated monitoring system. This allows human operators to monitor the surveillance area in real-time and make informed decisions.

7. Send Message: If suspicious activity is detected, the system can be configured to automatically send a message to a designated authority, such as security personnel. This message might include details about the detected activity and its location within the video frame.

8. Higher Authority: The message is sent to a designated authority, security personnel or law enforcement, who can then take appropriate action to investigate the situation.

9. Stored Video: The captured video footage can be archived on a storage device, such as a disk drive, for later retrieval and analysis. This archived video can serve as evidence in case of an incident or be used for forensic purposes.

E. comparison chart for kaggle datasets and real time video

When comparing the performance of 2D-CNN with the Kaggle dataset for human suspicious activity recognition, the accuracy rate is 90.88 higher than that of VGG16. ResNet50’s performance results in greater accuracy of 95.55 when in contrast to VGG16 and 2D-CNN using the Kaggle dataset. Using real-time video and 2D-CNN, the performance of human suspicious activity recognition yields a higher accuracy rate of 98.96 than VGG16. When comparing the accuracy of 99.01 for human suspicious activity recognition using ResNet50 to 2D- CNN and VGG16 with real-time video, the former performs better. An overall comparison chart using real-time datasets with VGG16, ResNet50, and 2D-CNN, and Kaggle datasets.

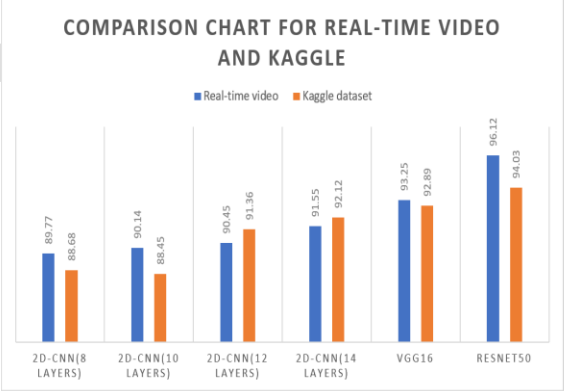


Fig. 7. Comparison chart

F. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) play a crucial role in suspicious activity detection within the context of video surveillance. CNNs are a class of deep learning models designed for image and video analysis, making them well- suited for tasks involving visual data. In the process of suspicious activity detection, CNNs can be employed for various purposes. Initially, during feature extraction, CNNs excel at automatically learning hierarchical representations of visual features such as shapes, textures, and patterns. This capability is particularly valuable when identifying complex and nuanced activities within crowded scenes, as CNNs can discern subtle visual cues that might indicate suspicious behavior. Moreover, CNNs can be integrated into the segmentation phase, helping to isolate relevant regions of interest in the video footage. Additionally, when coupled with recurrent neural networks (RNNs), CNNs can capture temporal dependencies, allowing the model to understand the evolution of activities over time. **Convolutional layer:** creates a feature map to predict the class probabilities for each feature by applying a filter that scans the whole image, few pixels at a time.

Pooling layer (down-sampling): scales down the amount of information the convolutional layer generated for each feature and maintains the most essential information (the process of the convolutional and pooling layers usually repeats several times).

Fully connected input layer: flattens the outputs generated by previous layers to turn them into a single vector that can be used as an input for the next layer.

Fully connected layer: Applies weights over the input generated by the feature analysis to predict an accurate label.

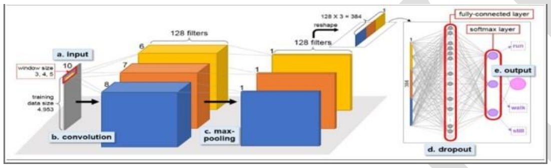


Fig. 8. Convolutional Neural Network General Architecture

G. *Sequence Diagram*

The sequence diagram illustrates a message flow between objects involved in extracting objects from a video. The mes- sages exchanged represent method calls between the objects. **1.Input Stimulus:** The diagram starts with an unnamed mes- sage representing the video and images being input into the system.

2. Frame Extraction: The Camera object transmits a message to the Extract Frames object, presumably requesting frame extraction from the video.

3. Object Extraction (Uncertain): There appears to be a message sent from the Classification object to the Extract Object block. The nature of this message is unclear without more context about the system. It's possible the Classification object is requesting the extraction of a specific object from each frame based on some prior analysis.

4. Iterative Extraction: The Extract Object block carries out object extraction on each frame, potentially multiple times per frame as indicated by the loop symbol. This could indicate the object extraction process is attempting to refine its extraction result, or it could be extracting different objects within a single frame.

5. Output (Uncertain): The purpose of the Output object is unclear from the diagram. It could represent the Classification block informing another system component concerning the classification results, or it could signify another stage in the object extraction process such as post-processing the classification output.

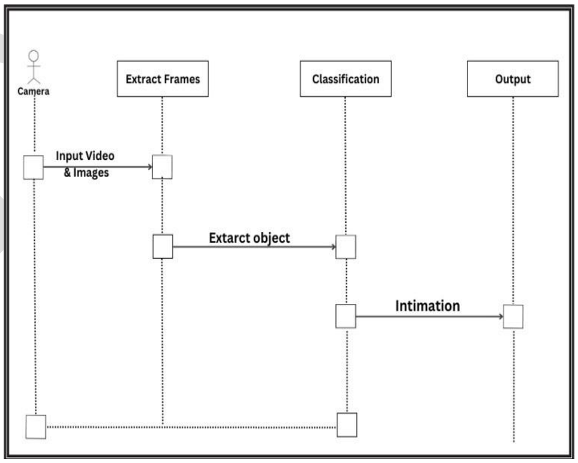


Fig. 9. Sequence Diagram *Resnet50*

Preprocessing: In the initial stage, the system captures the video stream from the surveillance camera. Salient motion frames are then selected from the video stream for further processing.

Training: The system is trained on a dataset of pre-classified normal activity videos during this stage. A lightweight CNN is employed to extract spatial features from the video data. These features are subsequently fed into a Long Short-Term Memory (LSTM) network, which is designed to learn the temporal patterns of normal activity within the video data.

Testing: During the testing phase, the system receives a new video stream from the surveillance camera. Similar to the train- ing phase, the system extracts spatial features from the video frames and feeds this data into the LSTM network. The LSTM network then classifies the activity in the current frame as normal or abnormal based on the knowledge acquired during the training phase. If the network detects abnormal activity, an alert is triggered, potentially notifying the authorities.

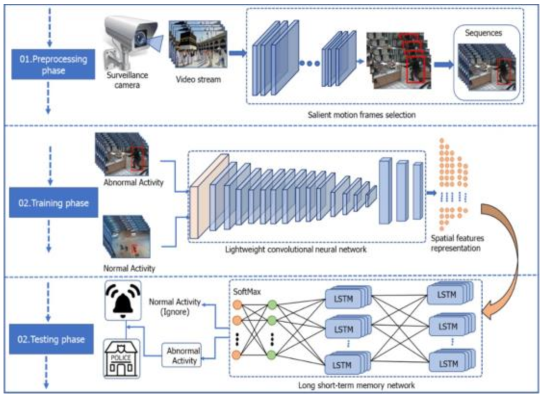


Fig. 10. Resnet50

H. *Computer Vision*

Computer vision plays a pivotal role in suspicious activity detection by providing advanced techniques for the automated analysis of visual information. In the context of video surveil- lance, computer vision algorithms are essential for extracting meaningful features from the captured footage. These features encompass various aspects such as object detection, tracking, and recognition of human activities. Computer vision enables the identification of anomalies or unusual patterns in video data, which can be indicative of suspicious behavior. Tech- niques like motion analysis, crowd density estimation, and object tracking contribute to the understanding of complex scenes and facilitate the recognition of activities that deviate from normal patterns. Moreover, computer vision algorithms can be integrated with machine learning models to enable the system to learn and adapt to evolving scenarios over time. By automating the analysis of visual information, computer vision significantly enhances the efficiency and accuracy of suspicious activity detection in video surveillance, providing valuable support to security and law enforcement efforts.

I. *Training and Testing*

YouTube videos, campus videos, KTH videos, and DCSASS datasets are the sources of the input videos. About 300 videos of both typical and suspicious behavior have been gathered by us. Preprocessing involves taking frames out of the video that was recorded. We employ the system's pre-trained VGG-16 model and its insights to solve the issue. Based on the requirements, the final layer of this model is eliminated, and the LSTM architecture is employed for classification. This is how our dataset is trained.

Fig. 11. testing

RESULTS AND CONCLUSION

The system demonstrated a suspicious module that pro- duced an accurate background without creating any artificial place paths or inefficient pixels. The AT module reduced the computational complexity for the ensuing motion detection phase by doing away with the need to examine the entire background region once a high-quality background model hadbeen generated.

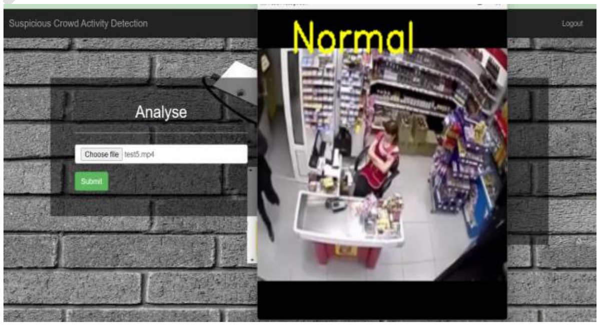


Fig. 12. Normal Activity

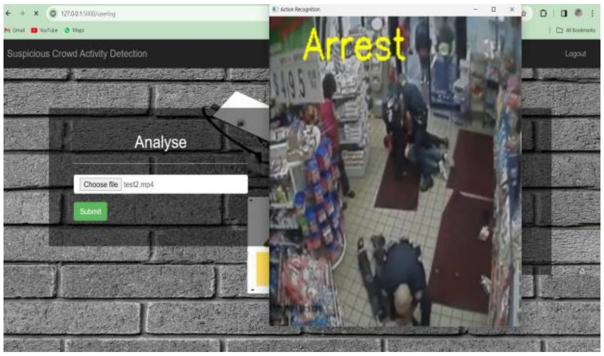


Fig. 13. Suspicious Activity

From the moving object mask, the suggested object extraction module identified the moving object pixels inside the triggered alert region. It also begins the process of creating a system for tracking suspicious people and analyzing their actions. At last, this algorithm functions for processing videos in real-time online with minimal computational overhead. The system has the potential to be utilized in the future with a highly accessible storage service and an advanced method of video capturing in surveillance areas.

REFERENCES

[1] Andini Fal Dessai, Prof. Shruti Pednekar.(5 May 2023). ”Surveillance- based Suspicious Activity Detection: Techniques, Application and Chal- lenges”. In 2023 International Journal of Creative Research Thoughts (IJCRT) (ISSN : 2320-2882).

[2] Indhumathi .J,Balasubramanian .M,Balasaigayathri .B.(February 8, 2023).”Real-Time Video based Human Suspicious Activity Recognition with Transfer Learning for Deep Learning”.In 2023 I.J. Image, Graphics and Signal Processing .

[3] Shashank Reddy Nallu, Vamshi Krishna Kunuru,Harshavardhan Reddy, Praveen H.(April 2023).”Suspicious Activity Detection from Surveillance Video using Deep Learning”. In 2023 JB Institute of Engineering and Technology (ISSN : 2349-6002).

[4] Digambar Kauthkar,Snehal Pingle.Vijay Bansode.Pooja Idalkanthe, prof. Sunita Vani.(6 June 2022).”Suspicious Human Activity and Fight Detection using Deep Learning”. In 2022 International Journal of Innovative Science and Research Technology (ISSN No:-2456-2165).

[5] Prof. Akshay Agarwal, Mr. Swapnil Galhate, Ms. Shipra N. Suvarna.(12 Mar 2020) ”Suspicious Activity Detection through CCTV” In 2020 International Journal for Research in Engineering Application and Management (IJREAM) (ISSN : 2454-9150).

[6] Rachana Gugale1,Abhiruchi Shendkar,Arisha Chamadia, Swati Patra, Deepali Ahir.(06 June 2020). ”Human Suspicious Activity Detection using Deep Learning” In 2020 International Research Journal of Engi- neering and Technology (IRJET) (e-ISSN: 2395-0056).

[7] J. INDHUMATHI , M. BALASUBRAMANIAN .(11, September 2022) ”Real Time Video Based Human Suspicious Activity Recognition Using Deep Learning” In 2022 Advances and Applications in Mathematical Sciences (Pages 6627-6650).

[8] Komal V Shivthare, Purvaja D Bhujbal, Akshada P Darekar, Prof. Yuvraj N N4.(4 April 2021) ”Suspicious Activity Detection Network For Video Survillance Using Machine Learning.” In 2021 INTERNA- TIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS (ISSN :2456-0774).

[9] S. A. Quadri, Komal S Katakdhond.(2022) ”Suspicious Activity De- tection Using Convolution Neural Network” In 2022 SECAB IET, Vijayapur, India.

[10] Esakky Selvi , Malaiyalathan Adimoolam , Govindharaju Karthi, Kan- dasamy Thinakaran ,Nagaiah Mohanan Balamurugan , Raju Kan- nadasan, Chitapong Wechtaisong ,Arfat Ahmad Khan 8.(16 December 2022).”Suspicious Actions Detection System Using Enhanced CNN and Surveillance Video” In 2022 MDPI

[11] YATHESHVAMSI NAIDU K , HARIBABU P .(2023) ”AI Suspicious Activity Detection using Human Pose Estimation” In 2023 Raghu Institute of Technology (ISSN(O)-2395-4396)

[12] Elizabeth Scaria, Aby Abahai T , Elizabeth Isaac .(2016). ”Suspicious Activity Detection in Surveillance Video using Discriminative Deep Belief Network” In 2016 International Journal of Control Theory and Applications (ISSN : 0974-5572)

[13] Nandini. G, Dr. B. Mathivanan, Nantha Bala. R. S, Poornima. P (2018). ”Suspicious human activity detection” In 2018 International Journal of Advance Research and Development

[14] Ahmed Mateen Buttar, Mahnoor Bano1 , Muhammad Azeem Akbar, Amerah Alabrah , Abdu H. Gumaei.(10 march 2023).”Toward trust- worthy human suspicious activity detection from surveillance videos using deep learning” In 2023 Software Engineering Department, LUT University.

[15] Monali Ahire, Devarshi Borse, Amey Chavan, Shubham Deshmukh, Favin Fernandes.(2022). ”Suspicious and Anomaly Detection” In 2022 Vishwakarma Institute of Technology