Real Time Suspicious Activity Detection by CCTV Footage Using Python and Deep Learning

[1] R.Kalpna, [2] CH. ARAVINDH KUMAR, [3] G. SAITEJA, [4] B. SHIVA SAI,[5] K. SANTHOSHINI
[1] Assistant Professor, Department of ECE, Sri Indu College of Engineering & Technology, Hyderabad.
[2,3,4,5] U.G. Scholar, Department of ECE, Sri Indu College of Engineering & Technology, Hyderabad.

Abstract: Video surveillance has become increasingly important in the modern world, especially with the rapid advancement of technologies such as artificial intelligence, machine learning, and deep learning. By integrating these technologies, various systems have been developed to identify and differentiate suspicious behaviors from live video streams. However, human behavior is highly unpredictable, making it challenging to determine whether an action is normal or suspicious. This study employs a deep learning-based approach to detect unusual or suspicious activities within an academic environment. When such behavior is identified, the system automatically sends alert notifications to the relevant authorities. The surveillance process is carried out by analyzing sequences of frames extracted from video footage. The methodology is divided into two main stages. In the first stage, meaningful features are extracted from the video frames. In the second stage, these extracted features are used by a classifier to determine whether the observed activity is normal or suspicious. This approach enhances the effectiveness of surveillance systems and enables timely intervention in critical situations.

Keyword: Suspicious Activity, Video Surveillance, Convolutional Neural Network(CNN), Visual Geometry Group(VGG-19)

I. INTRODUCTION

In our modern world, ensuring safety and security in public spaces has become increasingly complex. Human activity recognition is valuable in variety of scenarios, and anomaly detection in security systems is one of among them. With the increasing demand for security, surveillance cameras have been widely set up as the infrastructure for video analysis. Suspicious activity is any observed behaviour that indicates a person may be involved in a crime or is about to commit a crime. There are several effective algorithms to automatically detect human behavior in public spaces such as video surveillance for artificial intelligence, machine learning and deep learning. Artificial intelligence helps computers think like humans. An important component of machine learning is learning from training data and predicting future data. Today, there are GPU (Graphics Processing Unit) processors and large databases, so the concept of deep learning is used. The combination of computer vision and video surveillance will ensure public safety and security. Computer vision techniques include the following steps: environment modeling, motion detection, moving object classification, tracking, behavior understanding and interpretation, and data fusion from multiple cameras. This method requires a lot of work to extract features in different video sequences. Supervised and unsupervised classification methods. Supervised classification uses manually defined training data, while unsupervised classification is fully computer-driven and does not require human intervention.

Deep learning, a subset of artificial intelligence inspired by the structure and function of the human brain, has demonstrated remarkable prowess in various fields, including image and video analysis. Leveraging deep learning models, particularly CNN and recurrent neural networks (RNNs), holds immense potential in detecting anomalous behaviors and suspicious activities in surveillance footage. The fundamental principle behind using deep learning for suspicious activity detection is to train models on vast amounts of labeled data, allowing them to learn intricate patterns and deviations indicative of suspicious behavior.

A CNN can learn visual patterns directly from image pixels. The proposed system will use CCTV footage to monitor the behavior of people on campus and alert them when something suspicious happens. The main components of intelligent video surveillance are event detection and human behavior recognition. Automatic understanding of human behavior is a difficult task. Different areas of the campus should be monitored by video surveillance and various activities. Video footage from the campus was used for testing.

The developed system can take real-time videos from CCTV as an input, it then takes frames from the video and gives it to the CNN model. This CNN model takes a single frame as input, passes it through some operation to detect the occurrence of ‘Shoplifting’, ‘Robbery’ or ’Fighting’ in the store and produces a video with labelled frames as output. Each output frame is either annotated with ‘Shoplifting’, ‘Robbery’ or ’Fighting’ tags along with the probability.

VGG19 is a convolutional neural network (CNN) architecture that has been widely used in various computer vision tasks, including suspicious activity detection using deep learning. Originally developed by the VGG at the University of Oxford, VGG19 is composed of 19 layers, including convolutional layers followed by max-pooling layers and fully connected layers. VGG19 is a powerful architecture, its large number of parameters can make training and inference computationally expensive, especially in real-time applications. As such, optimization techniques and hardware acceleration may be necessary to achieve satisfactory performance.VGG-19 model for suspicious activity detection and highlights its potential for use in real-world surveillance systems.

Thus, It is very difficult to watch public places continuously, therefore an intelligent video surveillance is required that can monitor the human activities and categorize them as usual and unusual activities.

II. LITERATURE SURVEY

Ahmed Mateen Buttar et. al. [1] has explained about three models Convolutional Neural Network (CNN), GRU, and ConvLSTM model. These models are trained on the same dataset of 6 suspicious activities of humans that are: Running, Punching, Falling, Snatching, Kicking, and Shooting. The dataset consists of various videos related to each activity. Different deep learning techniques are applied in the proposed work: preprocessing, data annotation model training, and classification.

Phalguni Kadam et. al. [2] explained in their paper that the traditional method of monitoring cameras requires constant human intervention. Using Deep Learning and Image Processing, the proposed work aims to eliminate time and effort wasted on monitoring video surveillance cameras. Predicting human behavior is almost impossible. Deep Learning is used to detect suspicious and non-suspicious activity and to warn the user if any suspicious activity is detected. The proposed system strives for the detection of real-world suspicious activities such as burglaries, assaults etc. in surveillance videos.

Waqas Iqrar et. al. [3] in their work ,they used Convolution Neural Network (CNN) to extract spatial information along with a Long Short-Term Memory (LSTM) approach for the rapid and precise sequential tracking of an identified object. This CNN-LSTM technique not only lowers the model's complexity but also improves its accuracy which allows it to be executed in real-time.

Om M. Rajpurkar et al. [4] the goal of this paper is to identify suspicious activity for Surveillance and alert the shop owners when suspicious activity is detected. Electronic Article Surveillance (EAS) systems are widely used in today's retail stores, but this system is not capable enough as the shoplifters can easily remove the tag or label from the product.

Ujwalla Gawande et al. [5] worked on Person-focused dataset that includes various behaviors of students in an educational institution, such as cheating, theft of lab equipment, fights, and threatening situations. This dataset ensures consistent and standardized identification annotations for individuals, making it suitable for detection, tracking, and behavioral analysis of individuals. In addition, we have increased the detection accuracy through an improved architecture called YOLOv5 and introduced an efficient method for detecting global and local anomalous behaviors. This method extracts motion features that accurately describe the person's movement, speed, and direction. Accuracy 96.12%, with an error rate of 6.68% compared to existing methods.

Alavudeen Basha A et al. [6] the goal of the paper is Automatic tracking and detecting unusual movement's problems in closed circuit videos was resolved. Firstly, the videos are converted into frames. Then from the obtained frames, humans are detected from the video using a background subtraction method. Then the features are extracted using a convolutional neural network (CNN). The features thus extracted are fed to a Discriminative Deep Belief Network (DDBN). Labeled videos of some suspicious activities are also fed to the DDBN and their features are also extracted. Then the features extracted using Convolutional Neural Network (CNN) are compared against these features extracted from the labeled sample video of classified suspicious actions using a Discriminative Deep Belief Network (DDBN) and various suspicious activities are detected from the given videoand results shows increase accuracy of 90% for the proposed framework for classification.

Sushank Pawar et al. [7] explained in their paper that the system makes excellent use of AI technologies, such as machine learning and deep learning methods, to analyze a variety of data sources and identify unusual patterns that may indicate questionable behavior. The suggested system integrates various processing stages to provide robust and precise detection. First, information is gathered and preprocessed to extract pertinent elements from data from a variety of sources, including sensor networks, video surveillance, and social media feeds. Then, a broad range of AI models, including

anomaly detection algorithms, pattern recognition techniques, and behavior analysis tools, are used to analyze these features. To understand and generalize patterns, these models are trained on labeled data that includes both normal and questionable behaviors. The deployment of distributed computing resources is part of the solution, enabling quick and effective processing and analysis of massive datasets

III. METHODOLOGY

Detecting suspicious activities in real-world scenarios is a critical task for ensuring security and safety, prompting the utilization of advanced deep learning methodologies such as VGG19 and convolutional neural networks for suspicious activity detection involves several steps as shown in Figure 1 below:

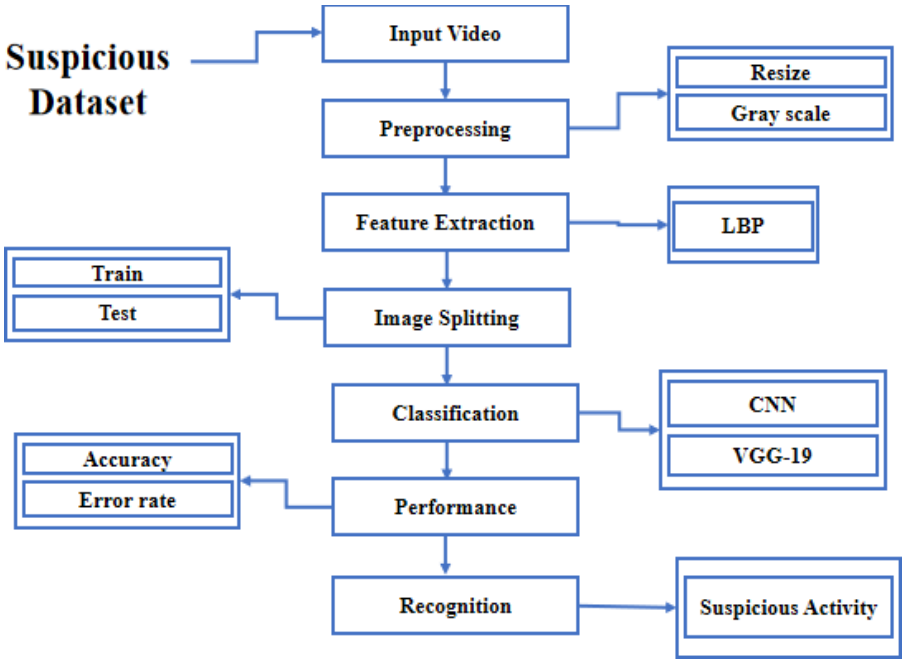

Figure 1:Flow Diagram

Input Video: The dataset contains the images in the form of '.jpg' or '.png'. In this step, the input image is read by using the imread () function. The input image is used to recognize the human activities. In this process, the tkinter file is used to dialogue box for selecting the input image.

Pre-Processing: In this process, the image has to be resized and convert the image into gray scale. To resize an image, call the resize () method on it, passing in a two-integer tuple argument representing the width and height of the resized image. The function doesn't modify the used image; it instead returns another Image with the new dimensions. Convert an Image to Grayscale in Python Using the Conversion Formula and the matplotlib Library. It can also convert an image to grayscale using the standard RGB to grayscale conversion formula that is,
imgGray = 0.2989 * R + 0.5870 * G + 0.1140 * B.

Feature Extraction: In this step, It is implemented or can extract the features from pre-processed image by using LBP. Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighbourhood of each pixel and considers the result as a binary number. LBPH (Local Binary Pattern Histogram) is a Face-Recognition algorithm it is used to recognize the face of a person. It is known for its performance and how it is able to recognize the face of a person from both front face and side face.

Image Splitting: During this process, data are needed so that learning can take place. In addition to the data required for training, test data are needed to evaluate the performance of the algorithm in order to see how well it works. Image splitting is the act of partitioning available data into two portions, usually for cross-validator purposes. One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance. Typically, when data set is separated into a training set and testing set, most of the data is used for training, and a smaller portion of the data is used for testing.

Classification: In this process, the deep learning algorithms are implemented such as VGG-19 and CNN.A **CNN** is a kind of network architecture for deep learning algorithms and is specifically used for image recognition and tasks that involve the processing of pixel data. **VGG** stands for Visual Geometry Group; it is a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. The "deep" refers to the number of layers with VGG-16 or VGG-19 consisting of 16 and 19 convolutional layers.

Performance: The Final Result will get generated based on the overall classification and prediction. The performance of this proposed approach is evaluated using some measures like, Accuracy of classifier refers to the ability of classifier. It predicts the class label correctly and the accuracy of the predictor refers to how well a given predictor can guess the value of predicted attribute for a new data.

IV. SYSTEM ARCHITECTURE

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. The Application Based On Suspicious Activity Detection.

This Is Software Based Application Figure 2 System Architecture

• Extracting frames from real time video. (i.e. video taken from CCTV)

• Pass the frame to trained CNN model.

• Push the predicted label for each frame to Q.

• Repeat step 3 for 'k' frames.

• Select the label with the highest probability corresponding to the mean of the last 'k' predictions

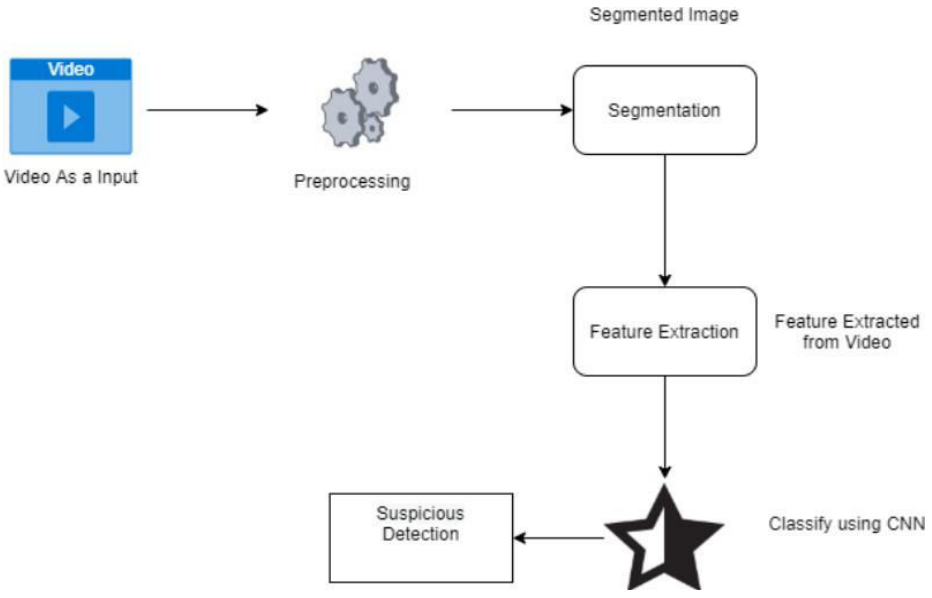

Figure 2: System Architecture

1. Preprocessing: Initial step involves preparing the video data for segmentation by removing noise, standardizing colors, and resizing frames for efficient processing.

2. Segmentation: Video frames are divided into meaningful segments or regions based on criteria like motion, color, or texture, enabling focused analysis on relevant portions.

3. Feature Extraction:Features such as shape, texture, and motion characteristics are extracted from segmented regions to represent key aspects of the video content.

4. Suspicious Detection:Extracted features are analyzed to identify anomalies or suspicious activities, such as sudden movements or unusual behavior, which may indicate potential security threats or safety concerns.

5. Classify using CNN:Utilizing Convolutional Neural Networks (CNNs), the extracted features are inputted into a classifier to categorize them as either normal or suspicious, allowing for automated real-time monitoring and response.

V. FUTURE SCOPE

Combining Convolutional Neural Networks (CNNs) with Visual Geometry Group (VGG) architecture for suspicious activity detection offers promising future prospects. This fusion enhances the model's ability to detect complex patterns in visual data, aiding in more accurate and reliable identification of suspicious behavior.

Moreover, advancements in CNNs and VGG architectures, coupled with ongoing research in computer vision and deep learning, will likely lead to even more robust and efficient models for security applications.

Integrating real-time monitoring systems and improving model interpretability could further extend its utility in various domains, including surveillance, fraud detection, and public safety.

VI. RESULTS

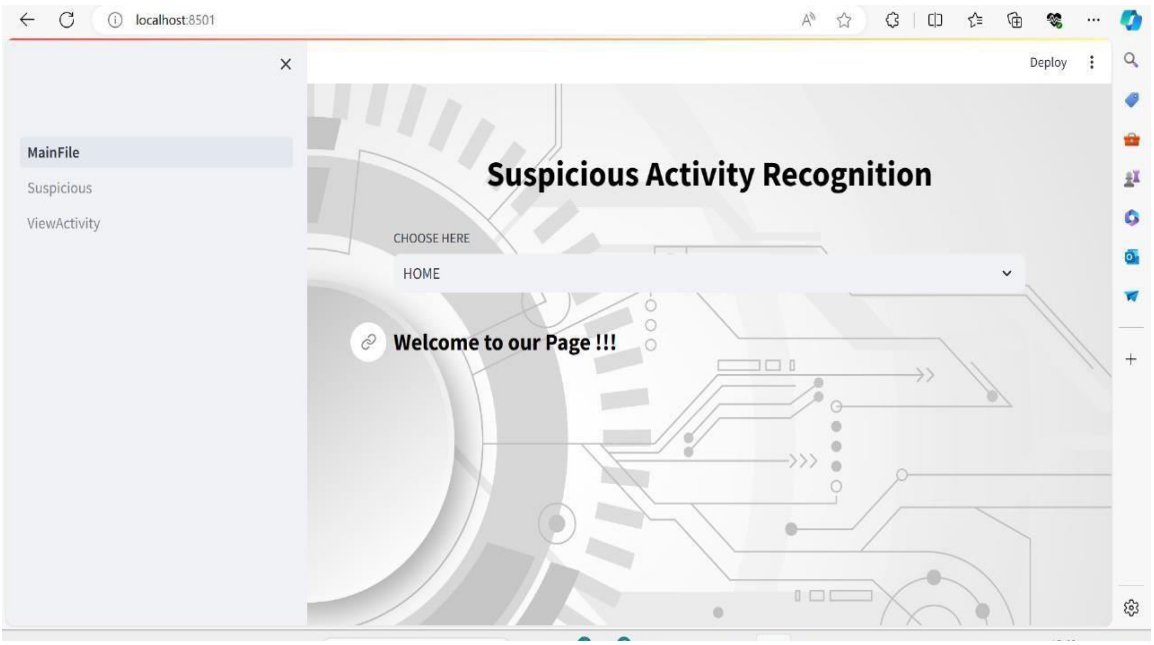

Figure 3: Home Page

In this section, the detailed designed and implementation of the system are presented.This part of the system gives the to convenient way to register and login himself.

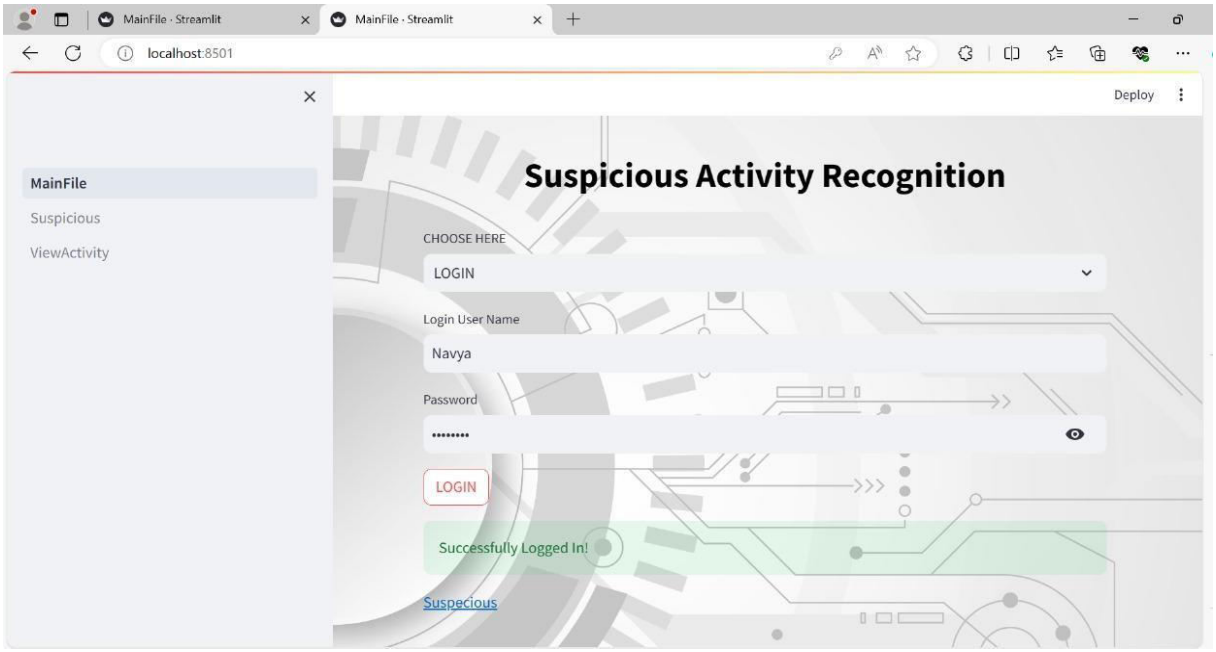

Figure 4: Login Page

Software Login interface :In this login interface we can login. If the username are correct then it will login and show the pop up of login successfully. It takes username and password for login .

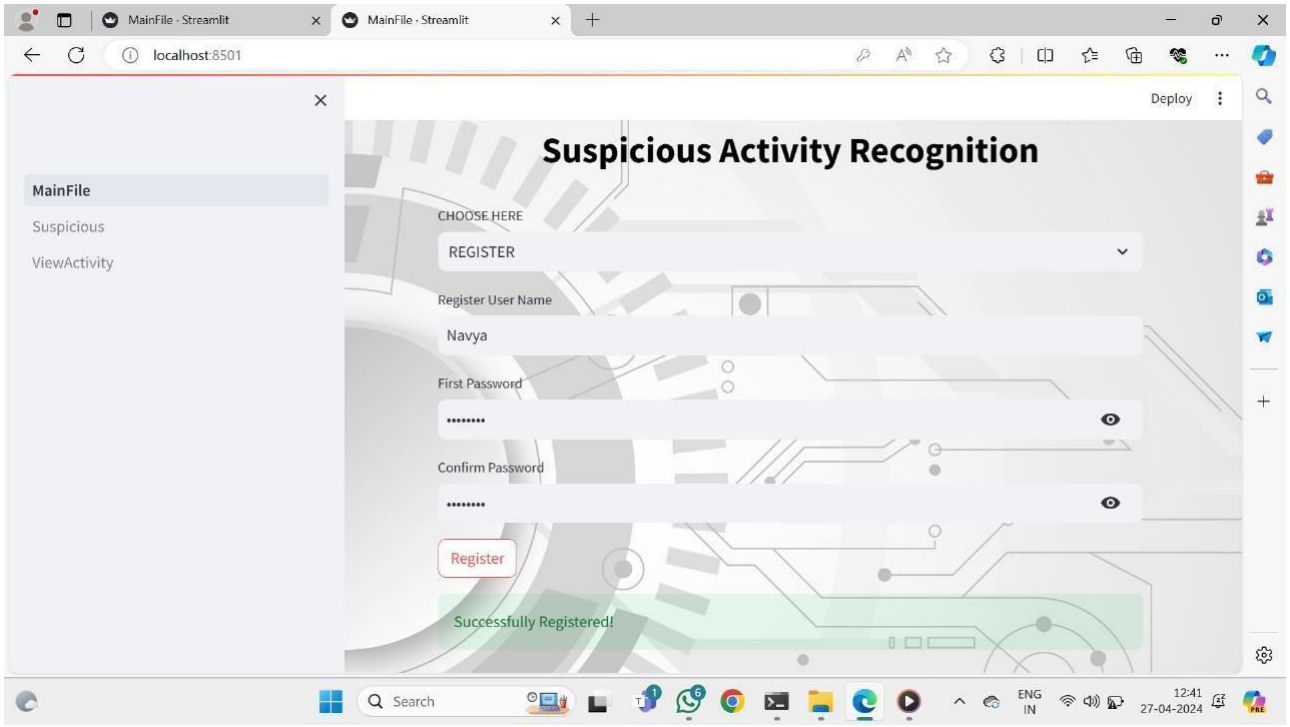

Figure 5: Registration Page

Software Registration Interface In this interface we can see that omkar can register themselves by using the name, address, email, phone number, user name, password .The email and password is the important credential to login the user again. After the registration successfully completed the registration process.

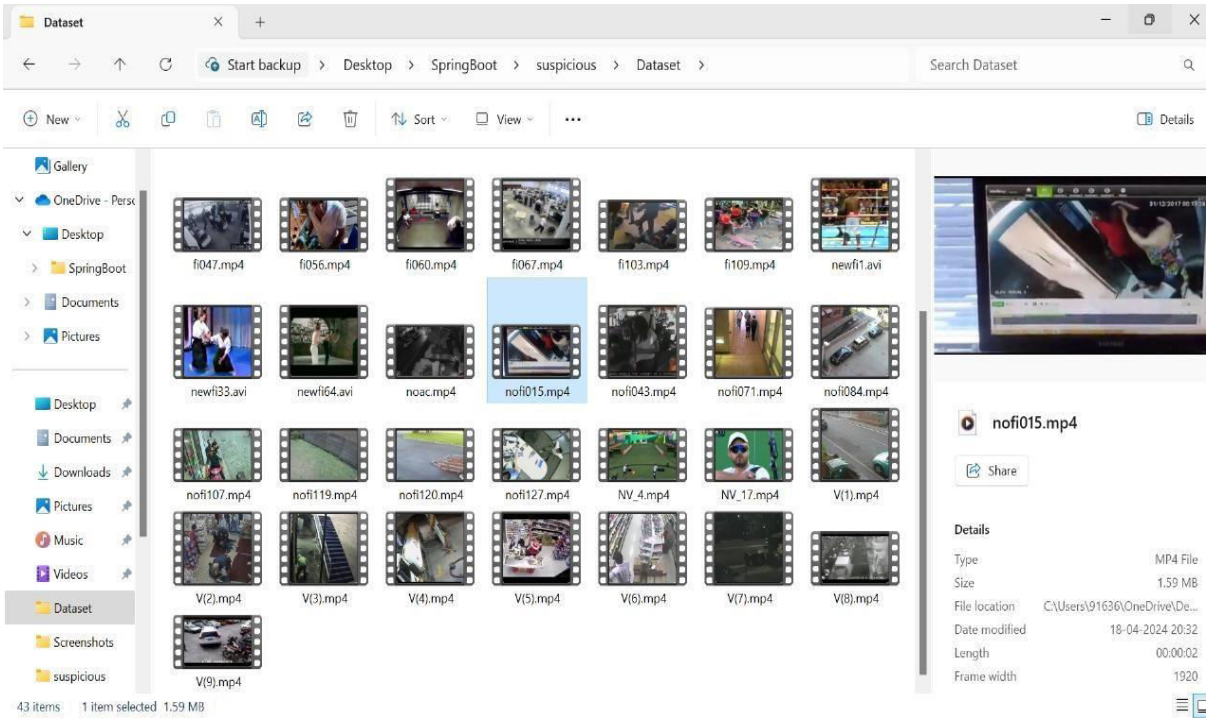

Figure 6:Input video's

Input Video: In this the stored video is taken as input video for detecting suspicious activity or not.

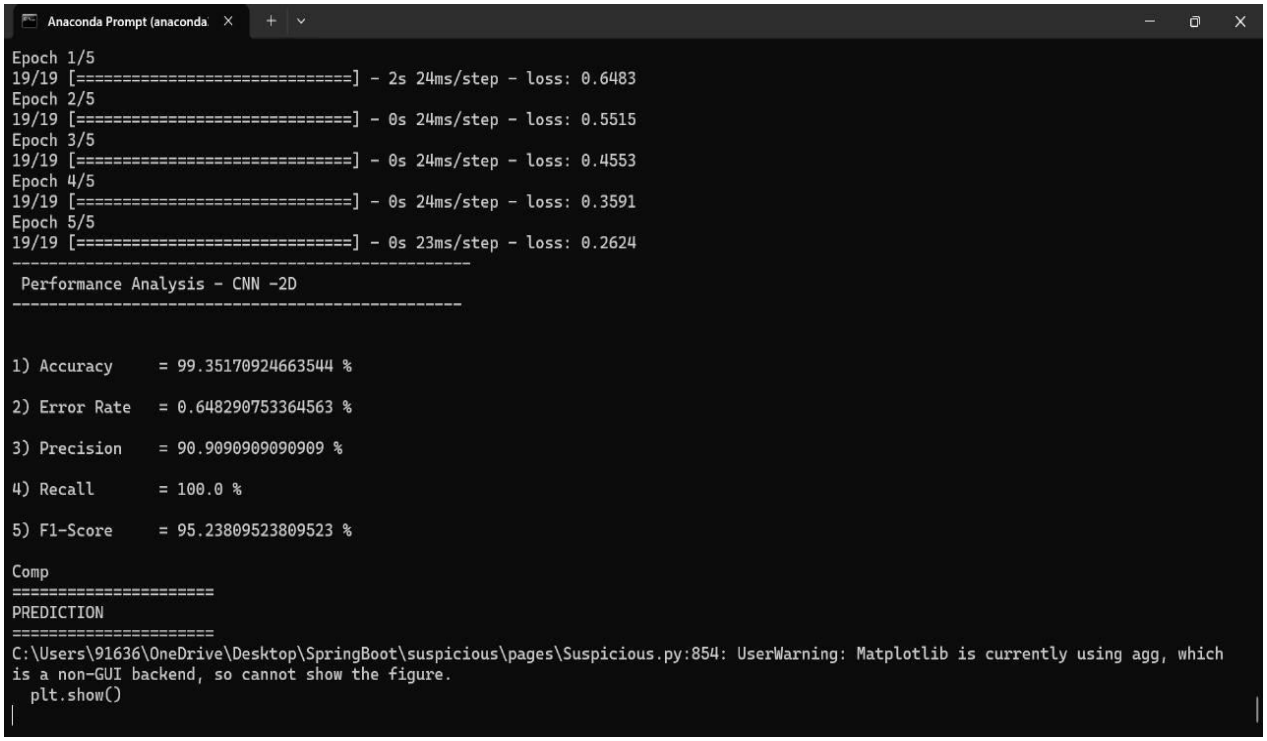

Figure 7: Result of Accuracy and Error rate

The **Final Result** will get generated based on the overall classification and prediction. The performance of this proposed approach is evaluated using some measures like, **Accuracy.**

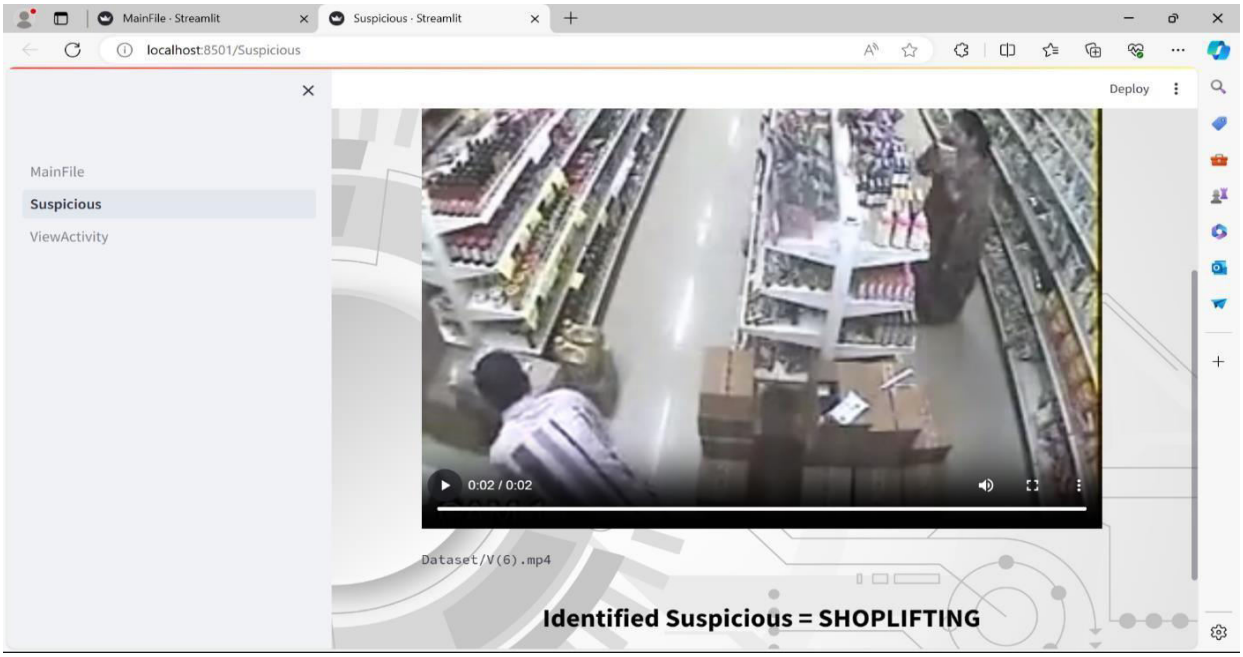

Figure 8: Identified Suspicious Activity

The result of suspicious activity detection system is shown in the above figure 8.

VII. CONCLUSION

The developed system can take real-time videos from CCTV or pre-recorded videos as an input, it then takes frames from the video and gives it to the CNN model. This CNN model takes a single frame as input, passes it through some operation to detect the occurrence of 'Shoplifting', 'Fighting', 'Accident' in the store and produces a video with labelled frames as output. Each output frame is either annotated with 'Shoplifting', 'Fighting', along with the probability. Finally, it is concluded that providing a system that determines customer behavior and detect suspicious activities without human intervention is a huge revolution in today's surveillance system.

REFERENCES

[1] Buttar, Ahmed Mateen, Mahnoor Bano, Muhammad Azeem Akbar, Amerah Alabrah, and Abdu H. Gumaei. "Toward trustworthy human suspicious activity detection from surveillance videos using deep learning." *Soft Computing* (2023): 1-13. o

[2] Kadam, Phalguni, Shweta Gawande, Akshita Thorat, and Rohini Mule. "Suspicious Activity Detection using Image Processing." *J. Sci. Technol* 6 (2021): 114-119. [3] " CNN-LSTM Based Smart Realtime Video Surveillance System ", 2022

[4] Rajpurkar, Om M., Siddesh S. Kamble, Jayram P. Nandagiri, and Anant V. Nimkar. "Alert generation on detection of suspicious activity using transfer learning." In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-7. IEEE, 2020.

[5] Gawande, Ujwalla, Kamal Hajari, and Yogesh Golhar. "Novel person detection and suspicious activity recognition using enhanced YOLOv5 and motion feature map." *Artificial Intelligence Review* 57, no. 2 (2024): 16.

[6]] Parthasarathy, P., and S. Vivekanandan. "Detection of suspicious human activity based on CNN-DBNN algorithm for video surveillance applications." In *2019 Innovations in Power and Advanced Computing Technologies (i-PACT)*, vol. 1, pp. 1-7. IEEE, 2019.

[7] Thombare, Puja, Vishal Gond, and V. R. Satpute. "Artificial intelligence for low level suspicious activity detection." In *Applications of Advanced Computing in Systems: Proceedings of International Conference on Advances in Systems, Control and Computing*, pp. 219-226. Springer Singapore, 2021.

[8] Parthasarathy, P., and S. Vivekanandan. "Detection of suspicious human activity based on CNN-DBNN algorithm for video surveillance applications." In *2019 Innovations in Power and Advanced Computing Technologies (i-PACT)*, vol. 1, pp. 1-7. IEEE, 2019.

[9] Quadri, S. A., and Komal S. Katakdhond. "Suspicious Activity Detection Using Convolution Neural Network." *Journal of Pharmaceutical Negative Results* (2022): 1235-1245.