Journal of Science and Technology
ISSN: 2456-5660 Volume 06,Special Issue 01, August 2021,

www.jst.org.in *DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp114-119*

Suspicious Activity Detection using Image Processing

Phalguni Kadam[1], Shweta Gawande[2], Akshita Thorat[3], Rohini Mule[4]
[1-4]*(Department of Electronics and Telecommunication Engineering, NBN Sinhgad School Of Engineering, SPPU, India)*
CorrespondingAuthor:kadamphalguni@sinhgad.edu, rohinimule@sinhgad.edu

To Cite this Article

Phalguni Kadam[1], Shweta Gawande[2], Akshita Thorat[3], Rohini Mule[4], "Suspicious Activity Detection using Image Processing ***", Journal of Science and Technology, Vol. 06, Special Issue 01, August 2021, pp114-119.***

Article Info
Received: 15.07.2021 *Revised: 24.07.2021* *Accepted:10.08.2021* *Published: 16.08.2021*

Abstract: *Nowadays, video surveillance is used almost everywhere for security. 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.*

Key Word*: Video surveillance, Deep learning*

I. Introduction

Human activities can be monitored in public places such as buses, railways, roads, banks, etc., to prevent crime through visual surveillance. Public spaces are constantly being monitored and intelligently monitored by video, detected and alerted.

In places where there is a risk of being robbed or shot, such as airports, train stations, and malls, our app can be used in surveillance. To train our system, we use deep learning and neural networks. This model will then be deployed as a mobile and desktop app which will take real time CCTV footage as input and send an alert on the administrator's device if some suspicious activity is detected. The challenge with anomaly detection is not so much in finding new techniques but rather in developing advanced algorithms that can detect a specific anomaly, such as a violence detector or a traffic accident detector. Removing these is a possibility if generative models are used. Nevertheless, these models are generally made on the basis of certain assumptions, such as uniform and independentnoise. This is due to the fact that even the most sophisticated algorithms can be influenced by image blur settings. Even so, such solutions do not work for other unusual events, so they have a limited application in practice. Human suspicious activity is related to identifying human body parts and possibly tracking their movements. It has a wide range of real-world uses, including gaming to AR/VR, to healthcare and gesture recognition[3,4]. The application of CNNs to video classification has received comparatively little attention as compared to image data domain. Finally, the main concept of anomaly detection is that unusual events are often imperceptible, and can only be detected by recognizing errors. This approach, though, tends to generate many false alarms due to changes in the environment over time (for example, at different times of the day).

II. Literature Survey

Eralda Nishani developed a real-time system meant to identify aggressive and violent behavior in real-time, thereby producing erratic and normal behavior patterns. Deep learning models was used to identify and quantify levels off in the video (CNN and RNN)[1].

Wagon Wan and Naimat Ullah Khan proposed a system that uses PIR sensors to detect human presence. They used Raspberry Pi to control motion detecting sensors and video cameras as well as to run remote sensing and surveillance equipment and also stream live video. The recordings from these systems are stored on the Pi,allowing for future playback. The surveillance system would be activated upon the detection of any movement[2].

Chen, Zhang, and Liu proposed a system that uses AMD algorithm to detect a suspicious person in a video and, once the user indicates one, will begin tracking the person. Complete detection of moving objects is accomplished with Advanced Motion Detection (AMD). To meet the rising demand for specialized monitoring equipment, a camera was also used in the monitoring room, where it was linked to the security monitoring equipment to emit alertmessages ifunusualactivitywasdetected[3].

Baole Ai, Yu Zhou, and Yao Yu used a hierarchical approach to identify suspicious activities by analyzing the differences in the motion characteristics between objects. The first step was to use a semantic approach to identify suspicious activities. After that, they used background subtraction to do object detection. After discovering evidence of human life, they further categorized the detected entities as living (human) or non-living (bag). The device was able to identify events as either normal or suspicious using motion features ortemporalinformation[4].

This system was developed by Zhe Cao that learns anomalies by combining videos of both normal and anomalous occurrences. In order to avoid having to annotate every video, they suggested using the Deep Multi-Instance Learning approach, which utilizes unlabelled training videos, in order to discover anomalies[5].

A deep neural network (CNN) for the new study's project, researchers Dongsung Lee, Hanguen Kim, and Sangwon Lee stated that it may be effective to teach a computer how various object appearances (background, tree, etc.) are connected to motion. A design that combines the two technologies was built using the same encoder. On 6real-world datasets, the experimental analysis demonstrated that the proposed approach was better than the mostcompetitivemethodscurrentlyavailable[6].

Tripathi, Rajesh, Jalal, Anand, and Agarwal proposed a system for detecting suspicious activities in images and videos using Convolutional Neural Networks (CNN). They investigated various CNN architectures and compared their accuracy. Their system architecture allowed them to process video footage from cameras in real time and detect suspicious activity. They also proposed potential future developments in this area[7].

Eksioglu, E. Decoupled MRI reconstruction technique based on a nonlocal block matching model, A new strategy for obtaining deep features, in particular, is proposed. The "Siamese Convolutional Neural Network" (S-CNN) was used to create the S-CNN technique discussed in this article, which delivers better results when identifying hyperspectral images. The goal of this strategy is to create a five-layered CNN that extracts deep characteristics first. The CNN's purpose is to resemble a nonlinear change function. The Siamese convolutional neural network is made up of two CNNs that were taught to look for features with low and high interclass variability. The S-CNN is a supervised technique that generates more discriminant features using the margin ranking loss function. For several hyperspectral datasets, the performance of a support vector machine (SVM) classifier was employed to compare our method to existing methods. It has been demonstrated that the method produces superior classification performance than conventional methods for feature extraction8,1.

S.Wang, Z.Su, L.Ying, X.Peng, S.Zhu, F.Liang, D.Feng, and D.Liang are among the original writers. Deep learning has the potential to improve magnetic resonance imaging (MRI). In the Ieee international conference International Symposium on Medical Science, a real-time violence detection system was developed that used deep learning to anticipate aggressive behavior in crowds or athletes. In a spark environment, frames were retrieved from real-time videos. If football violence is spotted, security personnel should be notified. To reduce the likelihood of violence occurring, the system detects the in -progress videos, notifies the security forces, and initiatesa counter measure. Was achieved using the VID dataset to identify violence in stadiums, which returned an accuracy of 94.5percent[3,9].

L. Xu, J. Ren, C. Liu, and J. Jia studied daily human activities, such as activities at home, work, caring, and helping, captured from videos. In the world of sports, training and practice rely on deep learning. RNN is utilized for the purpose of classification, and CNN is used for retrieving input features. They used the Inception v3 modelUCF101, as well as Activity-net and datasets, in this project[10].

As he discussed on stage, Yang looked into intelligent video surveillance for crowd analysis. This research looked at existing and advanced video surveillance methods and approaches, as well as several state-of-the-art deep learning algorithms and datasets[1,11].

It will be easy to trace down the individuals responsible if suspicious incidents captured on video footage are found. Humans can be seen in the videos using backdrop subtraction. The features were extracted by CNN and supplied to a DDBN (Discriminative Deep Belief Network). The suspicious footage is given to the DDBN, and its features are extracted as well[12].

III. Overview of the System

The framework relies on Raspberry Pi and is aimed to automatically detect suspicious activities and improve public safety and security. The suggested system will use footage from the camera to detect suspicious behavior in public locations and will send an alert if any suspicious activity is detected.

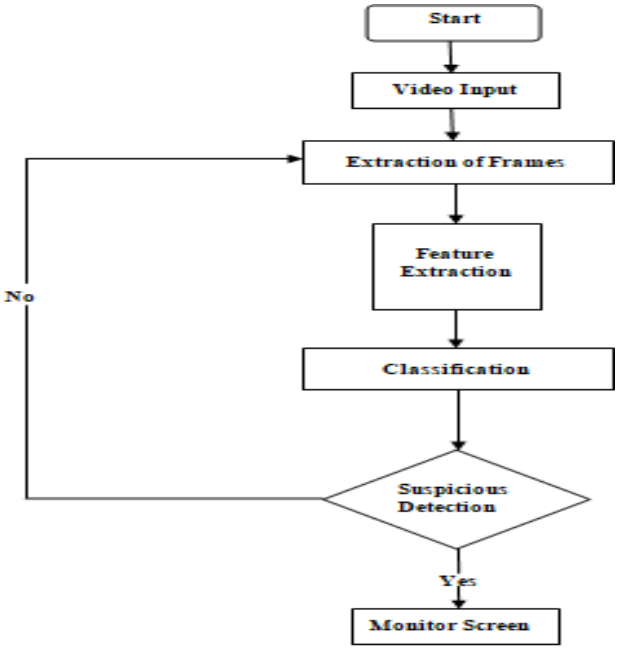
Proposed methodology:

Methodologies used in proposed system are

1. Image Processing: Image processing is the conversion of an image to a digital format and manipulation of that image in order to improve or extract relevant information from it. It's an information potential application in which the input is an image, like a video frame or a photograph, and the output is an image or image-related features. Typically, an image processing system treats images.These are their names:
2. Visualization: Focus on things you can't see.
3. Sharpening and restoration of images: To improve the appearance of the image.
4. Image retrieval: Find the image you're looking for.
5. Pattern measurement: Determines the distance between different objects in an image. Recognize the objects contained within an image using image recognition.
6. Deep learning: It's a neural network-based artificial intelligence function that extracts higher-level features from data and creates patterns for decision-making.

Flow Chart

Figure No 1 : Framework of Proposed Model



Working Principle

Preprocessing, feature extraction, and recognition phase are the primary components of this framework.

Preprocessing: The initial video is taken as input and fed into the system. During the preprocessing phase, the videois first converted into frames. The input video is divided into N continuous segments, with each video segment displaying a unique behavior pattern. Following segmentation, we use a Gaussian mixture model to separate background from foreground objects. It is employed in background subtraction. It is very important to remove unwanted noise from image to smooth the image.

Feature Extraction: The further step after noise removal is feature extraction. Extraction of features is the most important step in accurately recognizing various activities. Following video processing, features of successive frames are extracted. We consider parameters such as speed, movement, and direction when determining suspicious and non-suspicious activity.

CNN: Following feature extraction, CNN classification is used. For image classification, image recognition, object detection, and so on, CNN is the main source of information. Convolution, Linear, Pooling, and Fully Connected constitute CNN's four layers. The image is fed into the convolution layer.
Rectified to increase non-linearity, the linear unit layer applies an activation function to feature maps. The pooling layer then reduces the size of the input representation. It enables the detection of objects in images. The final layer, which is completely connected, is used to combine our features into attributes. This will improve the accuracy of class prediction[8,10].

When proposed system recognizes any suspicious activity it generates alarm to show suspicious activity by making its bounding box color as red and green if non-suspicious activity.
In this paper, the system is designed with the help of Raspberry Pi 3B+. This paper details the system's design, which includes a Raspberry Pi 3B+. In the design system, the features are satisfied on the Raspberry Pi 3B+, because it has a higher ethernet base and also Raspberry Pi 3B+ has a clock that runs at 1.4GHz.

Table No 1: Comparison of Raspberry Pi 3 and Raspberry Pi 3B+

Raspberry Pi 3	Raspberry Pi 3B+
• It has a processor. Broadcom BCM2837 Soc@1.2GHz	• It has a processor. Broadcom BCM2837 Soc@1.4GHz
• Ethernet: 100 base	• Ethernet: 1000 base
• PoE: NO	• PoE: YES
• WiFi: 802.11b/g/n	• WiFi: Dual-band 802.11 ac

III. Applications

One of the primary reasons for implementing a smart surveillance project is to provide authorities with real- time surveillance that can detect potential incidents and manage them as they occur. User-defined alerts and automatic unusual activity notifications are two sorts of alerts that can be generated by a smart surveillance system.

Motion detection, abandoned object alert, and behavioral alert are examples of user-defined alerts. This alert detects any moving object within a specific zone, such as unsecured baggage in an airport or a vehicle in a loading zone. It also detects crowds in checkout lines and alerts the store owner when the queue length exceeds a pre-determined threshold. In parking lots, this tool observes suspicious behavior, such as when someone in the line abruptly stops and opens several vehicles.

Automatic unusual activity alerts, in contrast to user-defined warnings, generate notifications when it identifies "behavior that deviates from the norm," such as detecting and recognizing cheating in the test room. Unexpected behavior identification is critical for efficient smart surveillance because all users cannot manually specify occurrences of interest.

IV. Result

This paper successfully demonstrated suspicious activity detection. The result of suspicious activity detection system is as shown in Figure 2 and Figure 3. The training phase's accuracy is 85.85 percent for the first ten epochs. After the detection, the output will be displayed on monitor screen with the use of python code.

Figure No 2 : Non-Suspicious Activity Detection

Figure No 3 : Suspicious Activity Detection

In the figures above, the green color indicates that the activity is not suspicious, while the red color indicates that the activity is suspicious. When a suspicious detection is detected, it will appear on the LCD screen and alert the user.

V. Conclusion

Almost everyone in today's world realizes the value of CCTV footage, yet in most cases, these recordings are used to aid in the investigation of a crime or occurrence. The proposed methodology offers the benefit of preventing crime from occurring in the first place. CCTV footage is being tracked and analyzed in real time. A system that processes real-time CCTV footage to detect any suspicious activity will help to improve security and reduce the need for human intervention. To detect real-world anomalies in surveillance videos, we proposed a deep learning approach. This is an attempt to exploit both normal and abnormal videos.

References

[1]. Nishani, E., & Çiço, B. (2017, June). Computer vision approaches based on deep learning and neural networks: Deepneural networks for video analysis of human pose estimation in 2017 6th Mediterranean Conference on Embedded Computing (MECO) (pp. 1-4) IEEE.

[2]. Naimat Ullah Khan, Wagon Wan (2018, July) A Review of Human Pose Estimation from Single Image"- 978-1-5386-5195-7/18/2018IEEE

[3]. Chen, Q., Zhang, C., Liu, W., & Wang, D. (2018, October). SHPD: Surveillance human pose dataset and performance evaluation for coarse-grained pose estimation in 2018 25th IEEE International Conference on Image Processing (ICIP) (pp. 4088-4092) IEEE.

[4]. Ai, B., Zhou, Y., Yu, Y., & Du, S. (2017, March). Human pose estimation using deep structure guided learning in 2017 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 1224-1231) IEEE.

[5]. Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fieldsin proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7291-7299).

[6]. Kim, H., Lee, S., Lee, D., Choi, S., Ju, J., & Myung, H. (2015). Real-time human pose estimation and gesture recognition from depth images using super pixels and SVM classifier Sensors, 15(6), 12410-12427.

[7]. Tripathi, R. K., Jalal, A. S., & Agrawal, S. C. (2018). Suspicious human activity recognition: a review. Artificial Intelligence Review, 50(2), 283-339.

[8]. Eksioglu, E. M. (2016). Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI. Journal of Mathematical Imaging and Vision, 56(3), 430-440.

[9]. Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., ... & Liang, D. (2016, April). Accelerating magnetic resonance imaging via deep learning in 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) (pp. 514-517).IEEE.

[10]. Xu, L., Ren, J. S., Liu, C., & Jia, J. (2014). Deep convolutional neural network for image deconvolution. advances in neuralinformation processing systems, 27, 1790-1798.

[11]. Yang, Y., Sun, J., Li, H., & Xu, Z. (2016, December). Deep ADMM-Net for compressive sensing MRI in proceedings of the30th international conference on neural information processing systems (pp. 10-18).

[12]. Zhan, Z., Cai, J. F., Guo, D., Liu, Y., Chen, Z., & Qu, X. (2015). Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction IEEE Transactions on biomedical engineering, 63(9), 1850-1861.