JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 10, ISSUE 03, 2023

"Securing Your Premises: A Comprehensive Framework for detecting unusual events by Video Analytics Software"

DR. NAGARAJA S R[1], MAJETY JYOTHIN MEHER VISHWANATH[2], NANDYALA MALLIKARJUNA REDDY[3], MADHUMITHA R[4], PADUCHURI AKANKSHA[5], VENKATA SANDEEP KUMAR REDDY[6]

[2,3,4,5,6]Student, Computer Engineering (Data Science), Presidency University, Bengaluru, Karnataka, India.

[1]Assistant Professor, Department of Computer Science & Engineering, Presidency University, Bengaluru, Karnataka, India.

nagarajasr@presidencyuniversity.in, jyothin.m.kumar@gmail.com, mallinandyala23@gmail.com, madhumitha.ramesh25@gmail.com, akanksha.paduchuri@gmail.com, cvsandeep0105@gmail.com

***ABSTRACT*—Anomaly detection is the process of identifying irregular, unexpected, unpredictable, or unusual events or things that aren't considered to be typical occurrences of those events or typical occurrences of those things in a dataset and hence vary from established patterns. Security is always a top consideration in any sector due to an increase in crime rates in busy events or possibly lonely areas.**

In our study, automated gun/weapon identification is implemented using a convolution neural network. The suggested method uses two different kinds of datasets. The photographs in one dataset were labelled beforehand, whereas those in the second dataset required manual labelling. Although both algorithms achieve high levels of accuracy, their practical use may require a trade-off between speed and accuracy.

***Key Words*: Automation, CNN, deep neural network, Twilio Tool, Haar Cascade Algorithms, Background subtract method.**

1.INTRODUCTION

We are all aware that certain events, such homicide and theft, occur frequently these days. Furthermore, the detector shouldn't be selective.[1] The creation of an AI-based detection system is the safest approach. It might not be sufficient to eliminate the human from the detection loop; the detector also has to be free from design restrictions. Domain-specific systems with personalized feature extractors could result in undesirable patterns. with the caveat that the training data distribution must be fair. In this sense, a connectionist machine learning system without a feature-based input space, such as a deep neural network, is vastly more egalitarian.[2] Support vector machines are used for precise item categorization. Another crucial factor is continuing to blend in. No specific person should be intentionally targeted via the warning system.[3] To notify individuals without causing alarm, a non-alarming audio-visual alert might be broadcasted in the vicinity of a social distance violation.[4] Numerous video sequences were examined by changing the window size. Software that is open-source must be used. This is necessary to build trust between society and the active monitoring system.[5] It uses embedded technology and a variety of sensors to keep an eye on its surroundings constantly for potential dangers like physical assaults, fraud, and theft that could endanger both the ATM and anybody around.

Our proposal is for an AI-powered active monitoring system that can detect breaches in social distance and send unidirectional visual or audio cues without being intrusive. A pre-trained deep-convolutional neural network (CNN) is used in the proposed technique to locate persons with bounding boxes in a monocular camera frame. Then, picture domain detections are converted into coordinates for a bird's-eye perspective of the actual environment. We recommend an AI-powered active monitoring system that can detect breaches in social distance and send unidirectional visual or audio cues to avoid becoming obtrusive. To avoid overcrowding, the system provides an advice in flow modulation signal if the social density exceeds a crucial level.

2.PROBLEM STATEMENT

Prior until now, manual security systems were used, and current camera technologies just record events, not make predictions. Devices link to an Internet gateway to communicate the sensor data they have collected. Although people may engage with

machines, the majority of work is now done automatically without human participation.

Traditional video surveillance systems rely too much on human oversight, which is exhausting and prone to mistakes. The need for constant monitoring is expanding, and failing to spot abnormalities in a timely manner runs the danger of having disastrous consequences. We were inspired to take on this project to show how this critically important activity can be automated and increase the efficiency of the surveillance system in order to reduce any risks associated with the non-detection of anomalies in time as and when they occur. This is due to the advancements in the availability of sophisticated video cameras, technologies for continuous streaming of video data, and Deep learning techniques.

However, occasionally due to a lack of security, some un-authorized users, hackers, or criminals take advantage of the opportunity to tamper the machine and steal the money by physically attacking automated teller machines. Other common ways that fraudsters abuse automated teller machines include cutting the safe, bombing, shoulder surfing, and ram riding. Numerous criminal activities target automated teller machines, including robberies, forced removal of the machines from the location, and assaults on the machines already there using explosives or tools for safe-breaking. In the majority of attack attempts, the criminal has had varied degrees of success. For automated teller machines, security mechanisms are occasionally added or invented to protect against this kind of fraud, but the criminals manage to take advantage of these security features.

3.LITERATURE SURVEY

[1] Unusual Event Detection in Low Resolution Video for enhancing ATM security

Authors: Sudhir Goswami, JyotiGoswami, Nagresh Kumar

Published in: 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 2015.

Methodology: To deal with low quality video, the proposed technique only employs near morphological operations with disk-like structural elements in the preprocessing stages. To distinguish foreground objects from moving backgrounds in a scene, it also employs rolling average background removal. By using a statistical trait called the standard deviation of moving objects, our proposed system may detect the existence of rare occurrences like crowding or fights in low quality videos.

Merits: It processes low quality footage quickly enough to be useful in surveillance systems for strengthening the security of ATMs, which still employ low resolution traditional cameras.

Demerits:. It does not employ a classifier and does not require initial system training.

[2] Abnormal activities detection for security purpose unattained bag and crowding detection by using image processing

Authors:SharayuSadashivPhule; Sharad D. Sawant

Published in: 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), 15-16 June 2017.

Methodology: This technology is intended to identify aberrant activity so that individuals may take appropriate precautions to avoid harmful events. One of the system's aberrant activities is abnormal crowd detection. In this case, the input will be a video, and the system's output will be the categorization of anomalous activity/object. This item might be a bag or a parcel with dimensions specified. For accurate object categorization, support vector machines are utilized.

Merits: Image processing using the support vector machine

Demerits: Slower processing

[3] A Novel Method to Enhance the Security of ATM using Biometrics

Authors: G. Renee Jebaline, S. Gomathi,

Published in: 2015 International Conference on Circuit, Power and Computing Technologies [ICCPCT] 978-1-4799- 7075-9/15/$31.00 ©2015 IEEE.

Methodology: In today's world, fraudsters can get PIN digits through a variety of means, necessitating the use of other security solutions such as biometrics. PIN numbers can be replaced with fingerprints, retina scans, and other biometric methods. Biometric technology currently analyses input photos to those in a database and, if they match, bank staff disburse cash. Our proposed idea, on the other hand, will totally replace PIN digits with biometric authentication. The machine will dispense cash after the comparison is satisfied.

Merits: As a result, the transaction's security is enhanced to a larger level.

Demerits: Processing for alternate techniques is slower if biometrics do not match.

[4] Robust Abnormal Event Recognition via Motion and Shape Analysis at ATM Installations

Authors: Vikas Tripathi, Durgaprasad Gangodkar, Vivek Latta, and Ankush Mittal

Journal of Electrical and Computer Engineering, Volume 2015.

Methodology: The motion history image (MHI) and Hu moments are used in the proposed method for extracting critical characteristics from videos. Support vector machines are used for classification and principal component analysis is utilized to minimize feature dimensionality. By altering the MHI window size, many video sequences were analyzed. With an average accuracy of 95.73%, the framework can distinguish between typical and aberrant acts such as theft, consumer damage from disagreement, or customer attacks.

Merits: A video-based framework that effectively detects anomalous actions at ATM installations and generates an alarm in the event of an unpleasant incident.

Demerits: low light processing is difficult

[5] Smart ATM Surveillance System

Authors:S.Shriram,SwastikB.Shetty,

Vishnuprasad P. Hegde , KCR Nisha, Dharmambal

Published in: 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT]

Methodology: This study focuses on an ATM surveillance system that uses embedded technology and a variety of sensors to enable continuous monitoring of its surroundings for potential dangers such as physical attacks, fraud, and theft that could endanger both the ATM and those around. The article also discusses the security measures that can be put in place to avoid such assaults through efficient surveillance. The many sorts of physical ATM assaults are discussed, as well as various strategies for detecting entry, taking preventive actions, and reporting authorities via the GSM network.

Merits: Using the GSM & other module make the Accuracy level high

Demerits: low efficiency in remote areas ATM

4.EXISTING METHOD

Most approaches for detecting abnormal occurrences in the literature rely on classifiers to identify the events rather than visual input. These classifiers need training data that is carefully monitored and learning time. Some methods have a significant computational cost and need human configuration initially in the automatic Video Analytic system. According to the research, in order to support a fully automated surveillance system, we require an algorithm that deals with detection of unusual events in the video.

5.PROPOSED METHOD

Our main contributions are: The detection of irregular, unexpected, unforeseen, odd occurrences or things that are not deemed to be regularly recurring events or regular items in a pattern or items that are present in a dataset and hence distinct from current patterns. Create a system that does not require human intervention to analyze data by watching continuous video.

• Ensure the system's analytical results are reliable and accurate.

Proposed system:

- The suggested system will significantly increase the security of the current system. Every action the individual makes is recorded by the camera, and those occurrences are saved as video clips, which may be used to spot unexpected activities and alert higher officials as needed.
- Image pre-processing utilizing the Background Subtraction technique.
- Image feature selection using the foreground subtraction technique.
- Image classification using the CNN algorithm.
- Detection of unusual events using the CNN and Haar Cascade algorithms.
- It is simple to recognize several uncommon events.

6.ALGORITHM: -

6(a). CONVOLUTION NEURAL NETWORK(CNN).

6(b)YOLO.

6(a)Convolution Neural Network:

Convolutional neural networks are one of the many feed-forward artificial neural networks., get

their connections between their layers from the visual cortex. Convolutional neural networks (CNNs), are used to evaluate the input data. They are used in a variety of disciplines, such as image recognition, video and image categorization, and natural language processing. The initial stage in extracting features from an input image is to apply convolution. Convolution allows for pixel associations to be preserved while collecting visual information from small data squares. As inputs, a filter or kernel and an image matrix are used in this process. The input image is subsequently processed by convolution layers, which employ filters (kernels) to generate output feature maps.

Here is an explanation of how algorithm works.

Modules:

1-Image Acquisition

2-Pre-processing of the Image

3-Image Segmentation

4-Extracting the Features

5-Classification Using CNN

Module 1-Image Acquisition:

The photographs are captured with a digital camera attached to a laptop. The acquired photos are subjected to further pre-processing.

Module 2- Pre-processing of the Image:

The photographs taken from the camera are pre-processed in order to improve their quality. Color transformation, noise reduction, histogram equalization, green masking, and other pre-processing processes may be used. In this case, we utilize color modification to improve the image's quality. Conversion of RGB picture to Grey scale and HSI to improve quality.

Module 3-Image Segmentation:

Image segmentation techniques include clustering, thresholding, neural network-based segmentation, and edge-based segmentation. For picture segmentation, we use the mean shift clustering technique in this implementation.

For converging to the maximum dense region, this approach employs the sliding window method. This approach uses numerous sliding windows to find the most dense region.

Algorithm for Mean Shift Clustering This method is mostly used to detect dense regions. Image Segmentation

Module 4- Extracting the Features:

A picture has several characteristics, the most important of which are color, texture, and form. We are looking at three features here: color histogram, texture that resembles color, form, and texture.

Typical CNN Architecture:

CNN's architecture, which seeks to resemble the neural network of the brain of a human being. A CNN's neurons are grouped in three dimensions and each evaluates a different aspect or characteristic of the input data/video and also the image.

To put it another way, each pair of neurons has been taught to recognize a certain feature of the image. In order to provide a final output that shows a vector of probability scores that indicate a

given feature's propensity to belong to a particular class, CNNs use layer predictions.

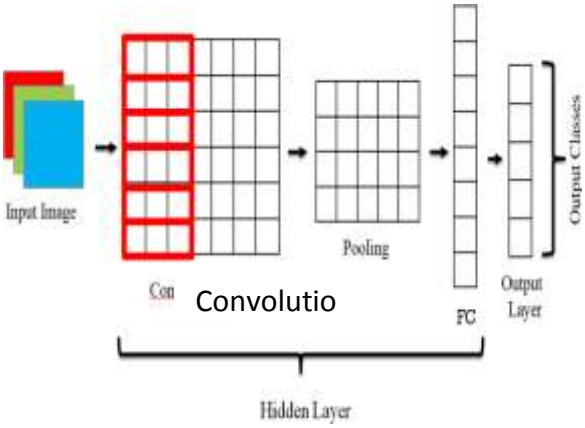


Fig 6(a)1.1-A CNN is composed of several kinds of layers:

Convolutional Layer:

Convolution layers are used to construct small portions of an image after it has been read by the computer as a series of pixels. Features or filters are used to describe these image parts or patches. The convolutional layer performs significantly better than comprehensive image matching situations at recognizing similarities by relaying these approximation feature matches in almost the same area in the two photographs. The fresh input photographs are compared to these filters, and the image is appropriately classified if they match.

Arrange the features and the image in a straight line, multiply each image pixel for pixel, add the pixels, then divide the feature's overall pixel count. We create a map and place the filter values where they belong. The feature will be moved to each additional point in the image, and its fit to each area will be evaluated. As a result, we will finally get a matrix.

ReLULayer

The rectified linear unit layer, or ReLU layer, substitutes any negative values in the filtered images with zero, guaranteeing that the values never go below zero. The activation function activates the node when the input value exceeds a certain threshold. If the input is negative, the function removes all negative values from the matrix and returns zero.

Pooling Layer

The pooling layer is in charge of lowering the image's size. This is accomplished by choosing a window size, entering the necessary stride, and putting the window above the filtered image. The highest values are then extracted from each window. This reduces the size of both the image and the matrix, which can then be mixed with other layers. The reduced matrix size is subsequently passed on to the fully linked layer.

Fully Connected Layer:

After the convolutional, ReLU, and pooling layers have been applied to the data, the next step is to stack these layers and then apply the fully connected layer for picture classification. These steps must be repeated if the input is not a 2x2 matrix. The completely connected layer is then used for categorization as the final step.

Output layer:

Produces the final probability for classifying the picture.

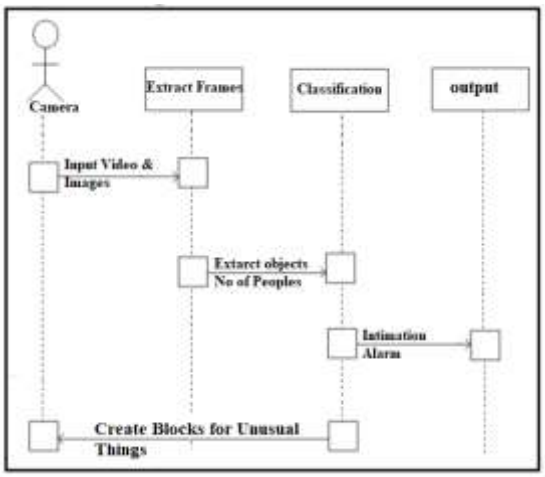
Sequence Diagram:

Sequence diagrams show the interactions between items in a chronological order. They display the classes and objects involved in a certain scenario, as well as the message exchange sequence required for the scenario to function. In the logical view of the system, sequence diagrams are frequently linked to use case realization's.

Individual processes or objects are depicted as vertical lines (lifelines) and the messages exchanged between them as horizontal arrows that follow their historical order in event diagrams or event scenarios, also known as sequence diagrams. This graphical depiction aids in the development of basic runtime scenarios.

The Sequence flowchart is depicted in Figure. The camera takes the input image provided by the user and, using the input image algorithm, extracts the number of objects and people in the given image. Furthermore, each item is categorized; if any of the objects is associated with an uncommon event, an automated Buzzer sound is made.

Fig 6(a)1.2-Sequence diagram



6(b).YOLO:

To distinguish between objects in real-time, the YOLO approach employs convolutional neural networks (CNN). "You Only Look Once," refers to the fact that by executing one forward run over a neural network for object detection, a single algorithm execution may anticipate the entire image.

Steps of Yolo algorithm:

Dataset collection: A dataset is a collection of several forms of data that is saved digitally. Data is the most important component of every Machine Learning project. Images are the most common type of data in a dataset. for tackling different Artificial Intelligence issues such as. Classification of images or videos.

Data annotation: Is the practice of marking data to illustrate the conclusion that your machine learning model should be able to predict. You are processing a dataset by labelling, tagging, transcribing, or otherwise adding the features that you want your machine learning system to learn to recognize.

Training: The model is trained using both the photos and the corresponding label file, which contains bounding box co-ordinates and class names for the objects in the images, once you've acquired them or trained the model from scratch using the label files. In this situation, classification and regression training take place simultaneously.

7.METHODOLOGY

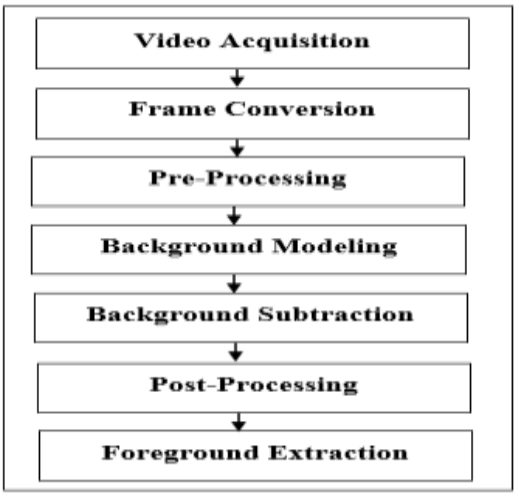


Fig7.1-Methodology Flowchart

Video acquisition: is the initial phase, which can be done with any of the technologies at hand, including a CCTV camera, a hand camera, a mobile camera, a USB camera, and more.

Converting the Frames: Once the video is captured, it is turned into frames so that appropriate processing may be done easily.

Pre-Processing: To decrease noise, pre-processing is performed to video frames. Smooth, dilate, erode, and median are some typical preprocessing methods.

Background Modeling: After pre-processing, backdrop modelling is used to generate either a static or dynamic ideal background based on environmental changes. Background modelling is an important stage in the system's capacity to perform numerous picture removal procedures. It is an essential part of all background removal methods. The literature describes a number of background modelling strategies that can be classified as recursive or non-recursive. The recursive approach is utilized in this system.

Background Subtraction: This is the most important stage in the background subtraction mechanism. Any major changes in the picture region from the backdrop model are discovered in this stage, and the pixels composing the regions experiencing change are tagged for future processing. To generate linked areas corresponding to the object, the connected component labelling algorithm is typically used.

Post-Processing: This is performed to enhance the findings. After background modelling and subtraction, there are several post-processing techniques available. The goal of these strategies is to enhance the foreground mask.

Foreground Extraction: Isolating the moving object from the frame is the method' final phase. The outcome of this stage is critical in determining how effective the background subtraction method is.

8.CONCLUSION

Finally, the ultimate purpose of the research is to detect and analyze unusual events. The ability to recognize anomalous patterns in pre-recorded and live data can provide valuable insights for industries such as security, banking, and healthcare. The Aim of the project is to create a sophisticated system that can recognize and analyze abnormal events in real time by utilizing cutting-edge technology such as machine learning, computer vision, and data analytics. This device can provide early warning signs of potentially hazardous situations, allowing for rapid intervention to avoid bad consequences. Overall, this research has the potential to have a significant impact on various enterprises as well as boost public safety by consistently and efficiently spotting and assessing unusual events.

9.FUTURE SCOPE

We could reinforce the project even further by putting more security in place in busy public areas like malls, train stations, and airports. By rapidly assessing CCTV data, the software might alert security personnel to the presence of firearms and help foil possible threats. analyzing camera footage from sporting events and sports broadcasts to look for any potential weapons that fans or spectators may be carrying. This could improve safety at sporting events and prevent any accidents.

10.REFERENCES

[1]SudhirGoswami, JyotiGoswami, Nagresh Kumar, "Unusual Event Detection in Low Resolution Video for enhancing ATM security", 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 2015.

[2]SharayuSadashivPhule ; Sharad D. Sawant, "Abnormal activities detection for security purpose unattainded bag and crowding detection by using image processing", 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), 15-16 June 2017.

[3]G. Renee Jebaline, S. Gomathi, "A Novel Method to Enhance the Security of ATM using Biometrics," 2015 International Conference on Circuit, Power and Computing Technologies [ICCPCT] 978-1-4799-7075-9/15/$31.00 ©2015 IEEE.

[4]Vikas Tripathi, Durgaprasad Gangodkar, Vivek Latta, and Ankush Mittal, "Robust Abnormal Event Recognition via Motion and ShapeAnalysis at ATM Installations", Journal of Electrical and Computer Engineering, Volume 2015.

[5]S.Shriram, Swastik B.Shetty, Vishnuprasad P. Hegde , KCR Nisha, Dharmambal V , " Smart ATM Surveillance System", 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT]