iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44825

www.ijraset.com

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

A Comprehensive Review on Advance Surveillance System

Prince Kumar Singh[1], Akarsh Singh[2], Shubham Ranjan[3], Aarjav Jain[4], Prof. Prajakta Dhamdhere[5]
[1, 2, 3, 4, 5]*Department of Information Technology, MIT ADT University Loni, Pune, Maharashtra*

Abstract: Advance Surveillance System has received growing attention due to the increasing demand for security and safety. It is capable of automatically analysing images, video, audio or other types of surveillance data with or without limited human intervention. The recent developments in sensor devices, computer vision, and machine learning have an important role in enabling such intelligent systems. This paper aims to provide an overview of smart surveillance systems and live CCTV detection. This paper also discusses the main processing steps in an Advance surveillance system: Object tracking, background- foreground segmentation, object detection and classification, and behavioural analysis

I. INTRODUCTION

Our goal is to design automated visual surveillance to reduce the burden on operators by including software in a surveillance system that can analyse video and manage risk content automatically.

Over the last few decades, remarkable infrastructure growths have been noticed in security-related issues throughout the world. So, with increased demand for Security, Video-based Surveillance has become an important area for research.

An Intelligent Video Surveillance system basically censored the performance, happenings, or changing information usually in phrases of human beings, cars or other gadgets from a distance via a few digital equipments (typically virtual cameras).

The scopes like prevention, detection, and intervention that have brought about the improvement of actual and steady video surveillance structures are able to shrewd video processing competencies.

In vast phrases, superior video-primarily based on total surveillance may be defined as a shrewd video processing method designed to help safety employees by means of supplying dependable actual-time signals and to aid green video evaluation and different investigations.

II. OVERVIEW

After ongoing discussions over various topics, we came together on this very topic that captured our interest, which was the peak increase in the number of security systems lagging usual safety detection of unwanted sources in their frames.

Also for the time being no such model has achieved satisfactory results in security systems. So of course, the matching sensor modality affects any potential methods for foreground-background segmentation. Using the BMC (Background Models Challenge) dataset, a paper recently compared 29 approaches. The top five effective approaches based on this experimental study include those that were suggested by everyone and also covered various sensing modalities in their survey article (such as audio, infrared, and thermal camera).

The majority of the background-foreground segmentation approaches that have been suggested use a single sensor modality, primarily visible cameras. Naturally, combining various sensor modalities would increase the system's robustness or streamline the segmentation processing.

For instance, the work of background-foreground segmentation is made simpler by merging visible camera and range data. By removing all range data outside of the observed area and using the visible/image data for fine segmentation, it is simple to filter out backdrop changes.

Collective points put forward by the four of us were brought together on a single conclusion, i.e. we need to work on a solution so as to provide a B2B and B2C solution at the minimum cost possible.

III. OBJECTIVE

1) To Design and develop a low-cost and time-efficient system.
2) It should decrease human labour as well as human contact.
3) Analyse millions of images and videos within minutes and augment human visual review tasks with artificial intelligence (AI).

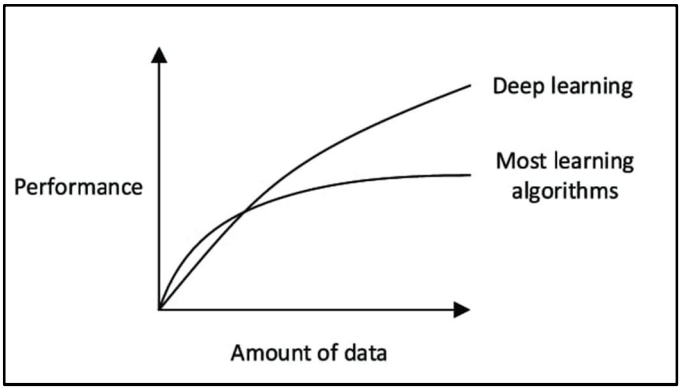
©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

IV. DATA PROCESSING TECHNIQUES

A. Foreground-Background Segmentation

The first stage in creating an intelligent surveillance system is to split the foreground and background. The idea is to differentiate the foreground (the object or moving object) from the background (the environment) (background). Many strategies for foreground-background segmentation, particularly for visible/video surveillance, have been presented. Several review papers that focus on foreground-background segmentation approaches are also available. Many research papers examined and provided a comprehensive description of most of the strategies that are accessible. The method that we have implemented for foreground-background segmentation is an Object detection module which segments the fore-ground objects with background noise by simply identifying them as shown below:



Background foreground segmentation is a famous subject matter in photograph evaluation today. Automatic programs for detection, class and evaluation, in photographs and films, are broadly used in lots of special industries. These programs regularly require a strong and suitable historical past foreground segmentation for best performance.

B. Object Detection And Classification

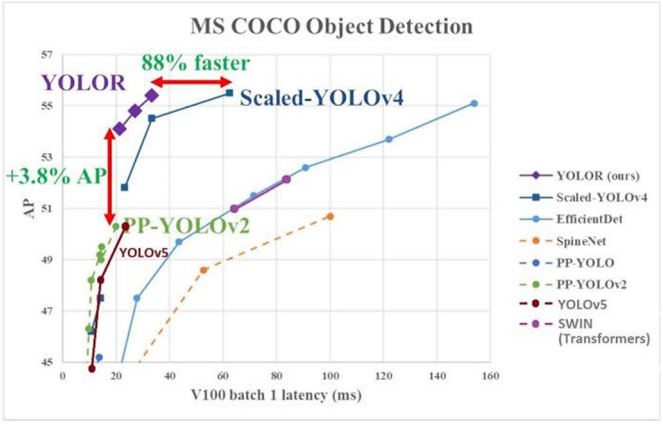
One essential element of an intelligent surveillance system is the ability to automatically detect and categorise objects, such as people and vehicles. Due to the huge range of possible appearances brought on by changing articulated position, clothes, lighting, and background, it is difficult for a machine (computer) to distinguish an item from a human. For people detection using a visual camera, numerous methods have been presented. One experiment revealed that HOG performed better at slower processing rates and higher picture resolutions, whereas the wavelet-based AdaBoost cascade technique performed better at faster processing speeds and closed real-time.

Recent benchmarking efforts have shown that FPDW has the best overall performance, however, MULTI FTR+MOTION is the best option if computational cost is not a factor. Additionally, a bottom-up, top-down detector for person detection based on lidar data was proposed. A bank of specialised classifiers for various height levels of individuals that together cast their votes into a continuous space is used by the bottom-up detector to learn a layered person model. The candidates are categorised using features that are computed in boosted volume tessellation voxels in the top-down stage. A people detection solution based on RGB Depth sensors—which offer both image and range data—was demonstrated while a map used 3D lidar to convert point data into depth images and perform people detection in 2D.

The majority of the proposed methods for classifying and detecting things focus solely on a small number of object types, such as people and cars. In reality, there are a lot of things that need to be taken into account in the real world, such as various animals and other things that could endanger security or safety.

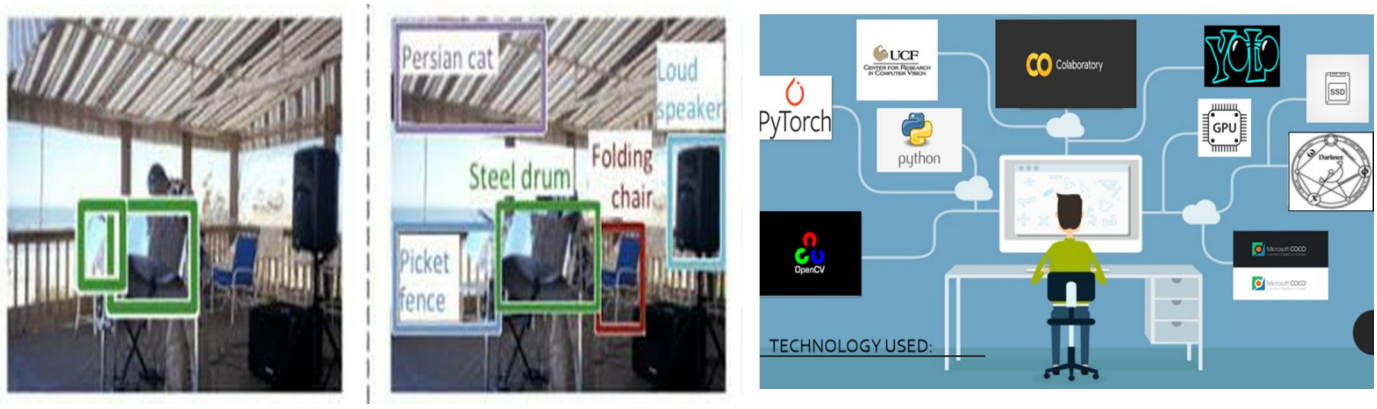
In this research, we found out that there are various sensors used for Object Detection but using these sensors in real-life is not feasible. So, we have used only one type of sensor which is already available in the form of CCTV network at various places. All the data that we get from these sensors or CCTVs will act as an input for a Deep Learning module which is specially trained over Microsoft COCOs Dataset which carries thousands of images of different class objects for predicting multiple objects present in a live frame.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



C. Object Tracking And Re- Identification

Typically, surveillance systems monitor the item in the spatial and temporal domain after object detection. Due to lighting variations, occlusion, clutter, sensor motion, and other problems, object tracking in realistic scenes is a difficult challenge. In recent years, many visual tracking algorithms (based on visible cameras) have been suggested. Model-based, appearance-based, contour and mesh-based, feature-based, and hybrid methods are the five categories into which object tracking techniques based on visual cameras can be divided. There are a number of review papers available that concentrate on visual tracking issues.



According to the results of the Visual Object Tracking challenge (VOT20I4), another attempt for assessing visual object trackers was put forth. The discriminative scale-space tracker (DSST) was recommended as the best tracker (combined accuracy and resilience). This tracker added a reliable scale estimate to the minimal output sum of squared errors (MOSSE) tracker. There have recently been several attempts to follow people using technologies other than visible cameras, such as radar, lidar, etc. For instance, real-time multi-person tracking using stereo range data. In addition to 2D photos, they also examined the range of data from stereo cameras. Using ground surveillance radar, they also created tools for automatically classifying targets, such as automobiles and pedestrians. Others utilised robots for surveillance using ultra-wideband (UWB) radars. They suggested 3D people tracking utilising a multi-target multi-hypothesis tracking approach based on lidar data. They used a bottom-up, top-down detector for people detection (explained in the previous section). Using probabilistic foreground modelling, multiple-person monitoring, and online re-identification, a paper suggested a method for real-time 3D people surveillance. The tracker module was also put to the test in actual outside settings, where there were numerous occlusions and a number of people who reappeared throughout the observation period.

There are two case scenarios while tracking an object, in the first one the object is in a dynamic state whereas in the second one the object is in a static state.. When the object is in the dynamic state then we can easily track them with a simple computer vision application but in the case of static objects it's very difficult to track them and on the other hand, re-identification is laborious due to feature extraction and feature matching algorithms.

That is why we have reversed the process of tracking and then re-identification to first identification and then tracking for this we split videos into picture frames and identified objects then bound them with bounding boxes while recording their coordinates. This is how we were able to track objects with better accuracy.

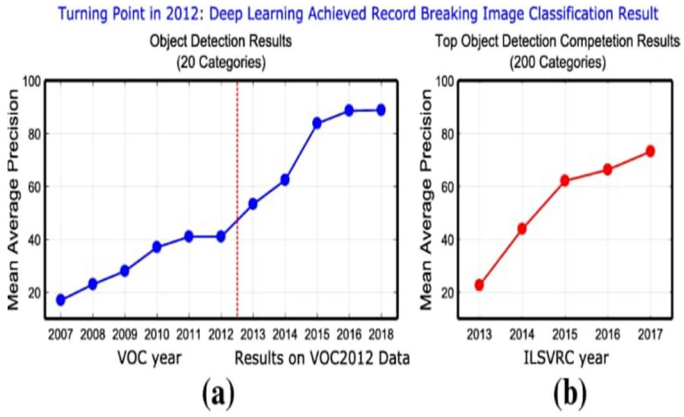
V. BEHAVIOURAL ANALYSIS

Automatic surveillance scene analysis is becoming more and more popular, not only for "object-level" (such as detecting and tracking) but also for "event level" analysis. Automated human behaviour analysis, group behaviour analysis, and event analysis are some areas of particular interest.

This subject has been covered in a few review publications. Human behaviour analysis may significantly improve security by speeding up the process of preventing undesirable events and spotting them even at the initial stage of suspicion. Even though it is important, analysing human behaviour is really difficult.

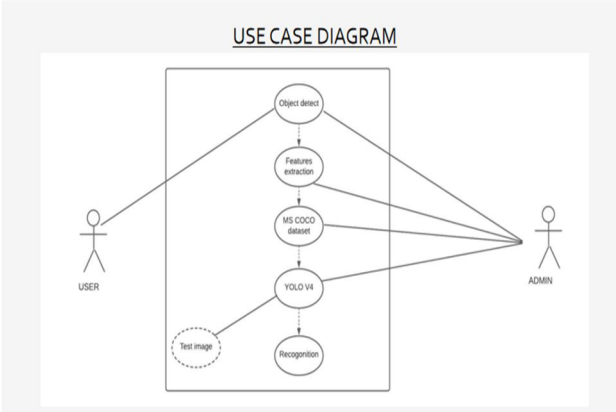
Classifying human behaviour is a fundamental element of human behaviour analysis. Human conduct has been categorised in a variety of ways. A simple classification of normal and abnormal was suggested in a paper. Add the categories of normal, odd, and abnormal to the categorisation.

The activities were previously categorised as positive, neutral, and negative activities in a paper. Humans can be observed as solitary beings, in groups, or in large masses. People fighting, being followed, walking together, and terrorist operations carried out in groups are a few examples of group activities. a technique for spotting five crowd behaviours in visual scenes: bottlenecks, fountainheads, lanes, arches, and blocking. In the context of visual monitoring of metro scenes utilising several cameras, they suggested an activity monitoring framework for identifying behaviours involving either isolated individuals, groups of people, or crowds.



There should be research into more sophisticated and realistic scenes. For instance, actual "fighting behaviour" may involve the use of weapons like knives or guns, reducing contact between the groups or individuals engaged in the fight. It's possible that these methods fall short of capturing this combative behaviour.

The size of weapons is usually very small so identifying them with good accuracy is also a challenge



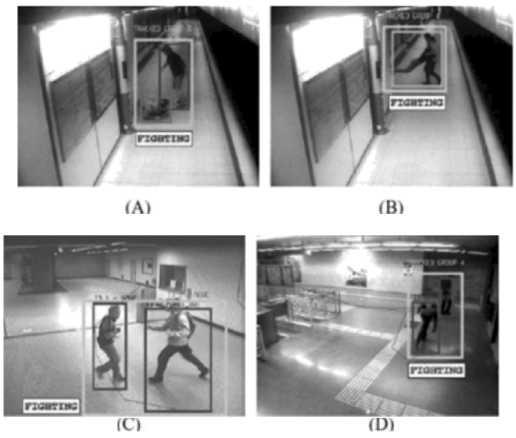


Fig. 7. Recognizing "fighting behavior" in a metro station [88]

VI. CONCLUSION AND FUTURE DIRECTION

A broad review of advanced surveillance systems has been provided in this paper. These intelligent systems hold promise for implementation in a variety of settings and uses.

Background-foreground segmentation, object identification and categorization, tracking, and behavioural analysis are some of the primary processing processes that have been the subject of numerous suggested methodologies.

Even though a number of encouraging findings have been made, additional research is required before actual adoption in settings with more complexity. For instance, new sensor modality combinations should be investigated while doing background-foreground segmentation in order to strengthen the system or streamline the processing.

Studies on behaviour analysis still focus on simplified scenarios, hence it is important to look into more realistic and complex scenarios. Researchers ought to think about looking into and creating advanced surveillance systems given the falling costs of processing technologies.

We were able to identify small objects in the frame but with less accuracy and the researched solution for this is instead of classifying the person and the small objects respectively we can train our model over a class for person, carrying that object.

REFERENCES

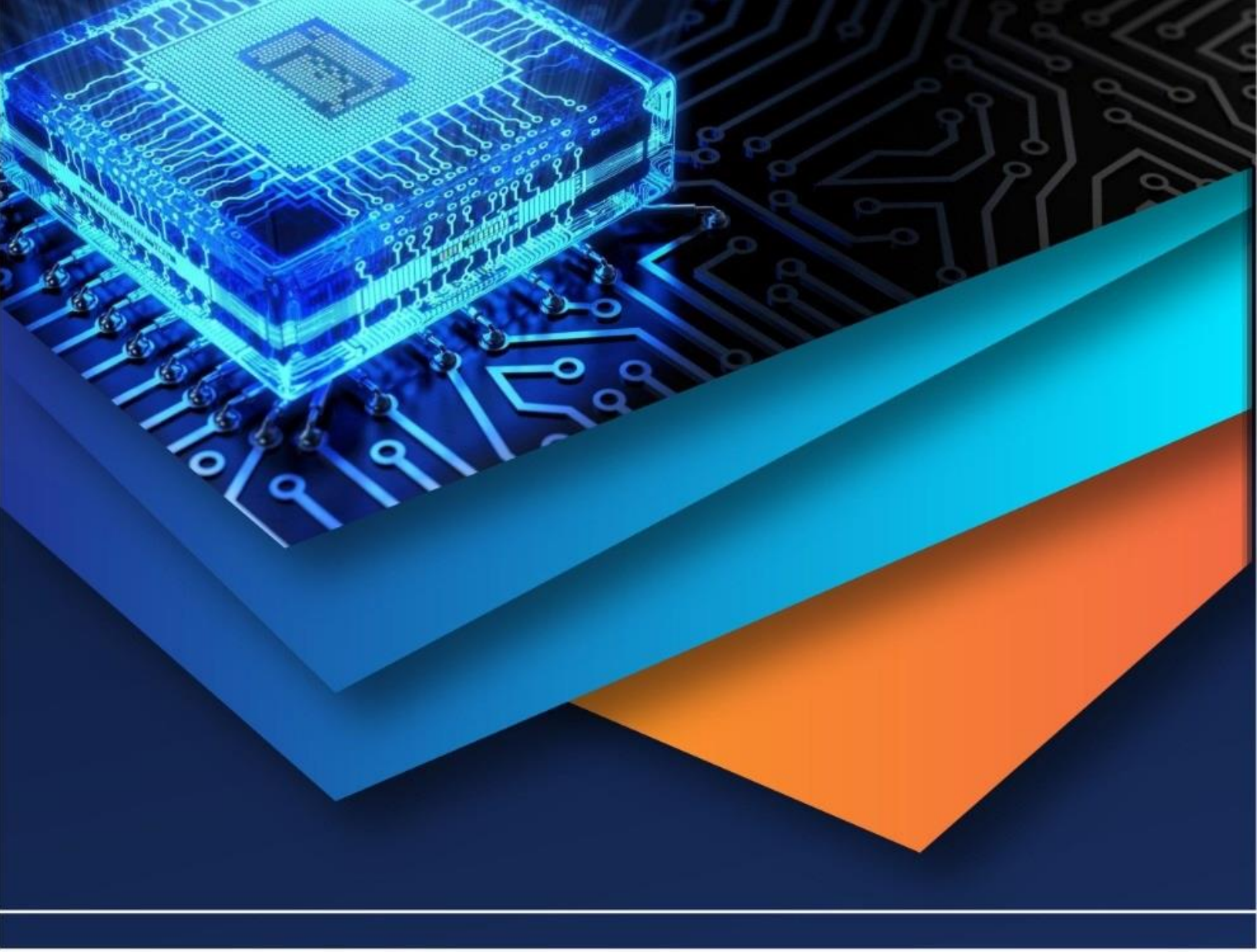
[1] T. Bouwmans, F. El Baf, and B. Vachon, Statistical background modelling for foreground detection: A survey, Handbook of Pattern Recognition and Computer Vision (Volume 4). Singapore: World Scientific, 2010

[2] A comprehensive review on intelligent surveillance systems Sutrisno Ibrahim* Electrical Engineering Department, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421 , Saudi Arabia

[3] N. Sulman, T. Sanocki, D. Goldgof and R. Kasturi, How effective is human video surveillance performance?, 19th International Conference on Pattern Recognition, Tampa, FL, 2008, pp.1-3R

[4] Liveness Detection for Face Recognition in Biometrics: A Review by Meenakshi Saini1, Dr Chander Kant2 1,2fDepartment Of Computer Science and Applications, Kurukshetra University, Kurukshetra, India)

[5] Microsoft COCO: Common Objects in ContextTsung-Yi Lin Michael Maire Serge Belongie Lubomir Bourdev Ross GirshickJames Hays Pietro Perona Deva Ramanan C. Lawrence Zitnick Piotr Dolla.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)