International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 3927 – 3935

# Detection and Identification of Unattended Objects in Video Surveillance using Machine Learning

**Mansi Santosh Jagdale[1], P. Venkata Rohit[2], Arnav Shivaji Patil[3], Swati Jagtap[4]**

*[1,2,3]Student, Department of Electronics & Telecommunication, Pimpri Chinchwad College of Engineering*

*[4]Faculty, Department of Electronics & Telecommunication, Pimpri Chinchwad College of Engineering*

[1]msj25399@gmail.com
[2]pvrohit100@gmail.com
[3]patilarnav62@gmail.com
[4]swati.jagtap@pccoepune.org

***Abstract***

*Video surveillance frameworks, acumen insight structures utilized in places for example railway stations, airports or various public places, can be helped to carry security to an upper level. The frameworks for video perception that are used for security reasons need advanced intellectual and robust technical ordinance. As the distress about the security across the world is soaring, it has raised the need to have installation of potential danger acknowledgment structures. This project depicts a framework which helps in recognizing the event of an individual leaving baggage unattended at a public place either deliberately or erroneously. Tracking of purposely left baggage is a significant issue as it obtrudes consequential security threats in crowded public transportation destinations in nations like India which has the biggest railway framework. The proposed system summons a caution whenever an abandoned baggage is encountered and also identifies time and date when the baggage was abandoned.*

***Keywords—*** *unattended, baggage, object, detection, machine learning, video surveillance.*

## I. Introduction

In the past few decades, automated and intelligent security monitoring systems have grabbed the limelight due to its ever-emerging requirements for similar frameworks. In crowded places such as bus stops, airports, railway stations, foyers, educational institutions and shopping areas, it is really difficult for the operators to manually spot baggage that has been abandoned. Baggage purposely left behind can lead significant security risks, typically in countries with wide-reaching rail systems such as India. The second aspect is to keep an eye on the luggage accidentally left behind. The manual tracking of the baggage is eventually laborious and need a lot of personnel. Generally, a small bunch of people observe multiple displays that rotationally show different views of many areas at an instant, this subsequently reduce efficiency and bomb threat rate. To achieve untimely detection of these intimidating remarks, security frameworks which are automated can-do justice to this to a great extent.

Consumers requirement for unattended object analyser is to identify suspicious and abandoned items within time of its abandonment process and ensure that threats are quickly detected. As a surveillance application, the automatic identification of left objects requires excessive-degree of precision and computational complexity. It is low enough to provide live-time performance.



ISSN: 2233-7853 IJFGCN

International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 3927 – 3935

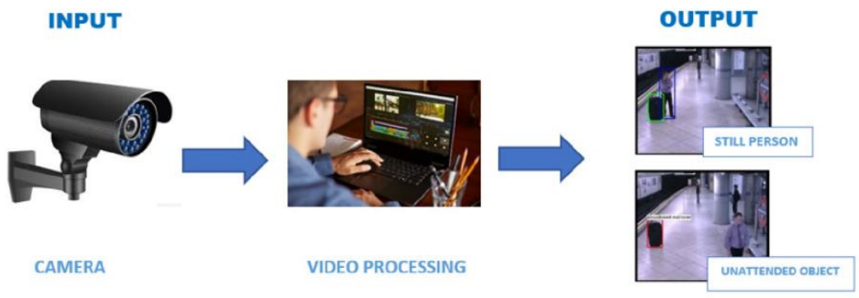



Fig. 1 Consumer Surveillance System

Our directives include the admission of four sub-events that help in describing the source of fascination. Once a bag is found, the structure will analyse its history to determine all the possible claimer(s). The possessor of the luggage is described as an individual who introduced the luggage in the frame before abandoning it.

Therefore, in a consumer surveillance framework, the detection of checked or left luggage and the detection of stolen objects is a spirited attribute of any monitoring system.

## II. Basic Concepts

*A. Machine Learning*: It is a field of Artificial Intelligence where computer programs can be used to solve real life problems by learning from past experiences and adapting based on the scenario. A computer can use a complex algorithm to scan data and make decisions accordingly. Machine learning is useful for analysing the vast amount of data available, advertising, credit, news, organization, fraud detection and many more.

B. *Video Processing*: Video processing is the field of signal processing. It can be used to detect and count objects in video files. A video is a single image or a sequence of images. Therefore, the algorithm aims to detect objects in images, and can also be converted to detect objects in videos.

C. *Image Processing*: It is the process of performing certain operations on an image to retrieve useful characteristics from it. Image is passed as an input and attributes corresponding to that image are obtained as output.

## III. LITERATURE REVIEW

The literature survey done for the topic "Detection and Identification of Unattended Objects in Video Surveillance using Machine Learning" is as follows:

Reference [1] proposes an automatic framework for the detection of abandoned objects which uses object tracking methodology in which the background is subtracted to extract the objects in foreground. The accuracy of this system is sufficient regarding the object detection but with conditions like glare and illumination on the object, the system fails to identify the abandoned object.

Reference [2] proposes a system that extracts the features of the objects which are unattended to identify the same using a base and static frame to retrieve the information from the frame. This system embraces an easy approach of utilizing the frames for information but this approach might give limited results in different scenarios.

Reference [3] mentions a method that detects the objects which are stationary using a hybrid model, the attributes of the object are extracted for classification of objects. Considering the computational complexity this system is not efficient as a scope of more accuracy was possible with inclusion of features such as association between human and baggage and various postures.



ISSN: 2233-7853 IJFGCN

Reference [4] implements a system that captures the events which lead to bags becoming abandoned using the logic based on temporal intervals to associate the events and actions. The approach used is highly intuitive and simple with good experimental results but it uses the reverse transversal method which is an old methodology.

Reference [5] uses the difference of dual background. It was modelled on statistical data of intensities of the pixels. This system has a high possibility to fail at detecting stationary objects as the time required for computation is comparatively very high and this system will only work if a static foreground mask is used.

Reference [6] is based on a two-level model. The first level is used to find passengers corresponding to unattended baggage and the second one reduces false alerts. The model being a two-stage system has less performance and isn't suitable for large scale deployment.

Reference [7] proposes to use a VGG model that can detect small objects. e.g., Knife detection. The model has a value of precision and recall as 58.5% and 54.3% respectively because of the high false positive and false negative.

## IV. PROPOSED MODEL

### *A. Single Shot Multibox Detector (SSD)*

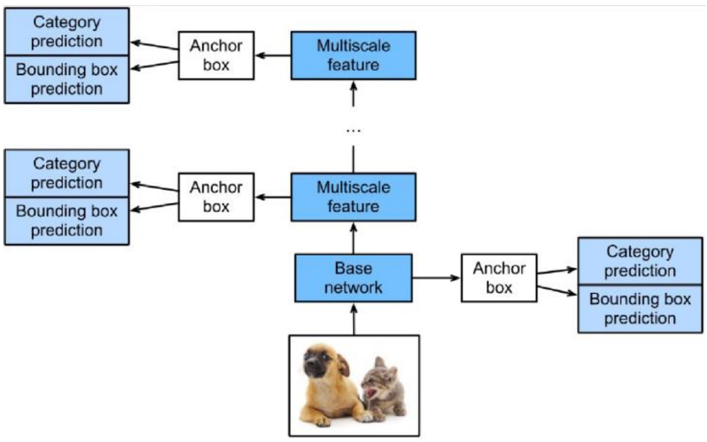



Fig. 2 SSD Architecture

SSD is a multiscale object detection model. The model's main components are multiple multiscale feature blocks connected in a series and a base network unit. Here the base network unit is employed to extract characteristics from the image. It is designed such that it outputs larger heights and widths and



ISSN: 2233-7853 IJFGCN

anchor boxes are retrieved from this feature map. Every multiscale feature unit decreases the size of the feature map received from the preceding layer. The nearer a multiscale feature unit to top, the fewer anchor boxes and smaller output feature map. Similarly, if a feature block is closer to top implies that it is easy to detect bigger objects. The model produces different anchor boxes of various sizes and is passed to a multiscale feature unit. Prediction of categories and bounding boxes are done by multiscale feature blocks.

B. *Block Diagram*

This section focuses on the implementation of the overall system with the help of a block diagram explaining the overall working of the project.

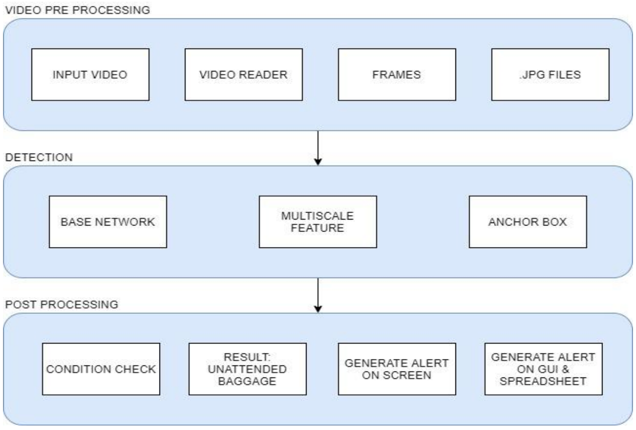



Fig. 3 Block Diagram

Description of the above block diagram:

*1. Video Pre-Processing:* A digital camera is used to capture the environment, and each image must be saved for processing, and the video is transmitted to the controller for further processing.

*2. Detection:* The base network module is used to retrieve characteristics from the photograph. The multiscale feature block reduces the size of the feature map received from the preceding layer. Anchor boxes permit a network to detect multiple objects of various sizes and overlapping objects. This eliminates the need to scan the entire photograph using a sliding window that computes a separate prediction for every viable location.

*3. Post Processing:* Display Device such as an LCD screen, is employed for authorized personnel to inspect the environment additionally to processor and take necessary action when needed. An Alarm is used to alert the concerned authorities in case of detection of any suspicious activity.

C. *Flowchart*



ISSN: 2233-7853 IJFGCN

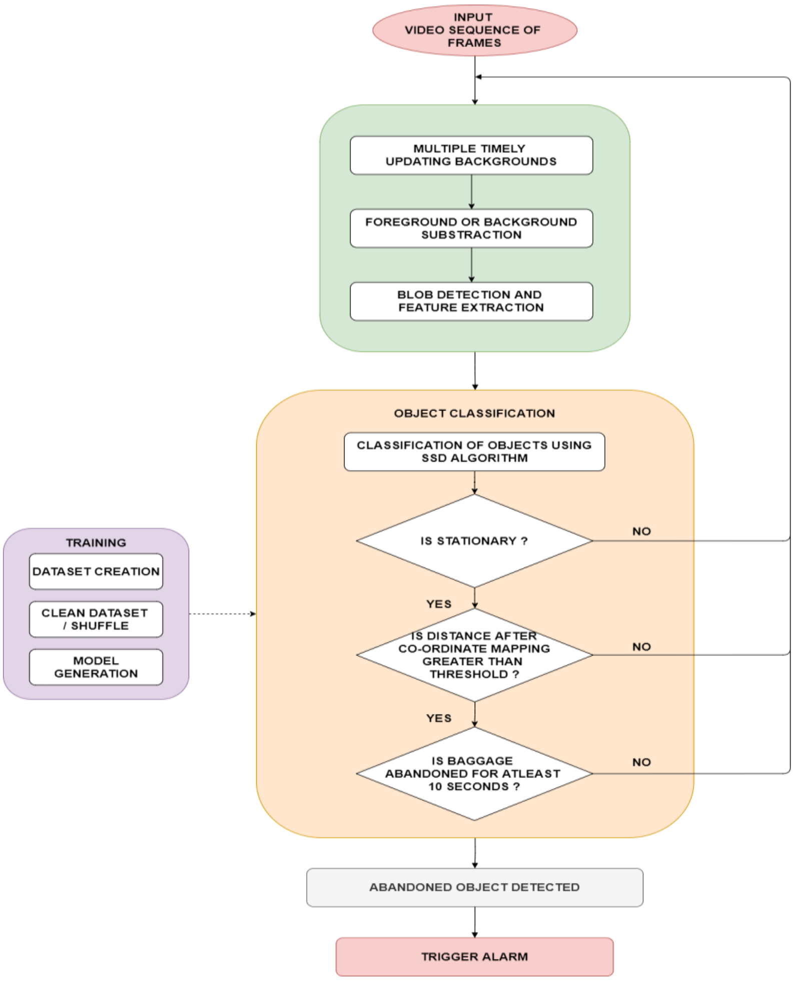


Fig. 4 Flowchart

Description of flowchart:

*1. Foreground detection and background subtraction:* Detection of moving objects from the distinction among the current frame and a reference frame, ("background image"). It provides essential clues for several applications in computer vision.

*2. Blob detection:* Blob detection strategies are geared toward detecting regions in a digital photograph that vary in properties, inclusive of brightness or color, in comparison to surrounding regions. The most common approach for blob detection is convolution.

*3. Feature extraction:* The process of retrieving useful characteristics from a raw image using machine learning algorithms like pattern recognition in image processing leads to better outcomes as it removes redundant data.

Features after extraction include:
*Detection anchor indices.* a tensor that contains the anchor indices of the detections and has shape [N].
*Detection boxes: [ymin, xmin, ymax, xmax].* a tensor that contains bounding box coordinates and has shape [N, 4].
*Detection classes.* a tensor that contains a detection class index from the label file and has shape [N].
*Detection multiclass score.* a tensor that contains class score distribution (including background) for detection boxes in the image including background class and has shape [1, N, 90].
*Detection scores.* a tensor that contains detection scores and has shape [N].
*Raw detection boxes:* A tensor that contains decoded detection boxes without Non-Max suppression. M is the number of raw detections and has shape [1, M, 4].



ISSN: 2233-7853 IJFGCN

*Raw detection scores.* a tensor that contains class score for raw detection boxes and has shape [1, M, 90]. M is the number of raw detections.
*Num detections:* A tf.int tensor contains a number of detections and has shape [N].

*Training model:* Training a model entails creation of a dataset to learn from, which is very well cleaned and shuffled. The column corresponds to a specific variable and the row represents an object of the dataset. A Dataset consisting of baggage images is built by scraping images from the web. Other sources of images include frames retrieved from a video recorded in railway station and airport.

*Object classification:* After the objects are detected in the frame different conditions are checked such as:
• Is the object(baggage) stationary?
• Is the distance between the object and its associated person(owner) greater than the threshold range?
Threshold range around the baggage is determined by relative sizes of the baggage image i.e., height, width. In this case threshold distance is (baggage_height+baggage_width)/3.
• Is the baggage abandoned for at least 10 seconds?
If all the above conditions are satisfied then an Alarm is generated alerting the concerned authorities.

## V. Results

Below table is for the testing using SSD MobileNet

TABLE I
SSD MobileNet Result Table

| **Test Case No.** | **True positive** | **False positive** | **TP(Acc)** | **FP(Acc)** | **Precision** |
|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 0 | 1 |
| 2 | 1 | 0 | 2 | 0 | 1 |
| 3 | 1 | 0 | 3 | 0 | 1 |
| 4 | 1 | 0 | 4 | 0 | 1 |
| 5 | 0 | 1 | 4 | 1 | 0.8 |
| 6 | 1 | 0 | 5 | 1 | 0.83 |
| 7 | 1 | 0 | 6 | 1 | 0.85 |
| 8 | 0 | 1 | 6 | 2 | 0.75 |
| 9 | 1 | 0 | 7 | 2 | 0.77 |
| 10 | 1 | 0 | 8 | 2 | 0.8 |

Precision=[TP(Acc)/TP(Acc)+FP(Acc)] Accuracy= [Σ Precision /10] *100 = 88%

Following table is for the testing using SSD EfficientDet

TABLE II
SSD EfficientDet Result Table

| **Test Case No.** | **True positive** | **False positive** | **TP(Acc)** | **FP(Acc)** | **Precision** |
|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 0 | 1 |
| 2 | 1 | 0 | 2 | 0 | 1 |
| 3 | 1 | 0 | 3 | 0 | 1 |
| 4 | 1 | 0 | 4 | 0 | 1 |
| 5 | 1 | 0 | 5 | 0 | 1 |
| 6 | 0 | 1 | 5 | 1 | 0.83 |



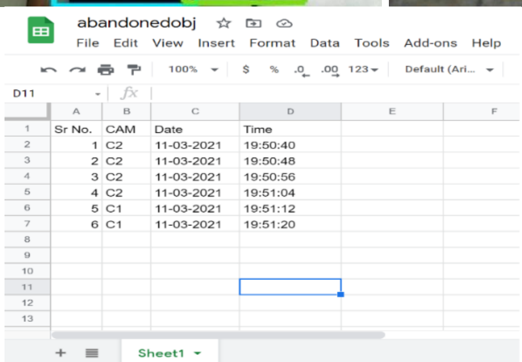
ISSN: 2233-7853 IJFGCN

| | | | | | |
|---|---|---|---|---|---|
| 7 | 1 | 0 | 6 | 1 | 0.85 |
| 8 | 1 | 0 | 7 | 1 | 0.87 |
| 9 | 1 | 0 | 8 | 1 | 0.88 |
| 10 | 1 | 0 | 9 | 1 | 0.90 |

Precision=[TP(Acc)/TP(Acc)+FP(Acc)] Accuracy= [Σ Precision /10] *100 = 93.3%

From the above tables, the true positive value is returned when the model is successful in predicting the required objects. A false positive value is returned when the model is unsuccessful in predicting the required objects. Precision is the ability of a model to identify only the required objects. It is the percentage of correct prediction. The following detection results were obtained during running a test case:






| Sr No. | CAM | Date | Time |
|---|---|---|---|
| 1 | C2 | 11-03-2021 | 19:50:40 |
| 2 | C2 | 11-03-2021 | 19:50:48 |
| 3 | C2 | 11-03-2021 | 19:50:56 |
| 4 | C2 | 11-03-2021 | 19:51:04 |
| 5 | C1 | 11-03-2021 | 19:51:12 |
| 6 | C1 | 11-03-2021 | 19:51:20 |

Fig. 5 Attended baggage

Fig. 6 Unattended Baggage

Fig. 7 Google Sheets view

A. *Frontend (GUI):* Simple GUI application is created to fetch data from server (Google Sheets) and display it in a well formatted manner. This application also gives an audio alert when an abandoned object is detected. PyQt5 and pyttsx3 libraries are used for GUI and audio effects respectively. Refer Fig. 8 GUI view.



ISSN: 2233-7853 IJFGCN

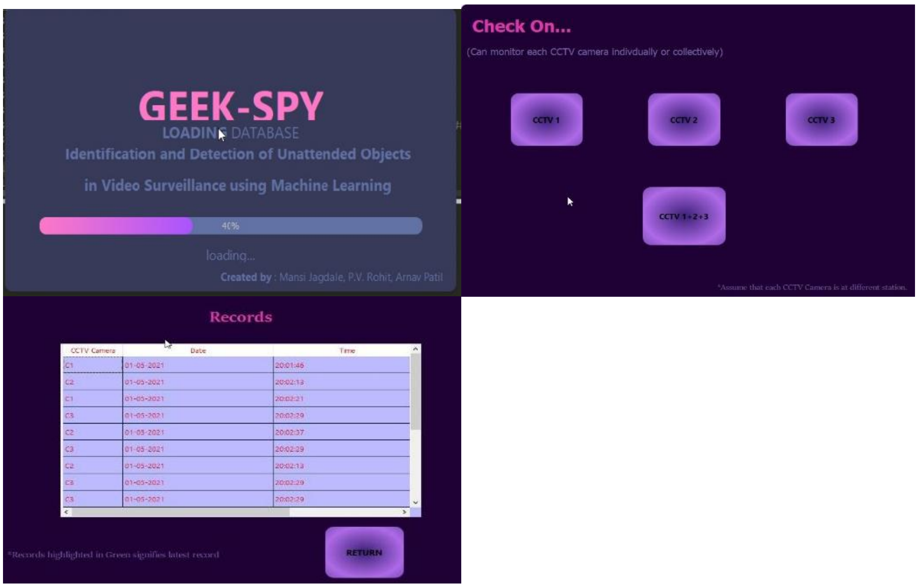



Fig. 8 GUI view

B. *Backend (Server):* Google Sheets is used as a server and data is continuously transmitted after a fixed time period. Gspread library which is a python API for interacting with Google sheets is used to send data such as Time, Date and corresponding camera id. Refer Fig. 7 Google Sheets view.

## VI. Conclusion

The project aims to identify and detect unattended bags and by using SSD MobileNet we were able to achieve 88% accuracy hence we considered using SSD EfficientDet and based on experimental results we are able to detect objects with accuracy of 93.3% and identify the objects individually with exact location of the abandoned object. The prediction is accurate enough and lies within the limits of experimental error. The system utilizes a static frame as a base frame to extract object information from frames which is a simple and easy approach.

## References

[1] Neelam Dwivedia, Dushyant Kumar Singh, Dharmender Singh Kushwaha: An Approach for Unattended Object Detection through Contour Formation using Background Subtraction. In: Third International Conference on Computing and Network Communications (CoCoNet'19)

[2] Pankit Arora, Akshath Dhar, Tushar Singh, Anju Mishra: Abandoned Object Identification and Detection System for Railways in India. In: IEEE International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT),2017.

[3] Lakhan H. Jadhav, Bashir Ahmed F. Momin: Detection and identification of unattended/removed objects in video surveillance. In: IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2016.

[4] M. Bhargava, Chia-Chih Chen, M. S. Ryoo & J. K. Aggarwal: Detection of Abandoned Objects in Crowded Environments. In: IEEE Conference on Advanced Video and Signal Based Surveillance, 2007.

[5] Wahyono, Student Member, Alexander Filonenko, Kang-Hyun Jo: Unattended Object Identification for Intelligent Surveillance Systems Using Sequence of Dual Background Difference. In: IEEE Transactions on Industrial Informatics, 2016.



ISSN: 2233-7853 IJFGCN

[6] Nam Trung Pham, Karianto Leman, Jie Zhang, Isaac Pek: Two-Stage Unattended Object Detection Method with Proposals. In: IEEE 2nd International Conference on Signal and Image Processing, 2017
[7] Yurika Fujii, Minoru Fukumi, Shin-ichi , Momoyo: Detection of Dangerous Objects by Pan-tilt Camera. In: International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 2019
[8] Intel. (2017, September 08). Unattended Baggage Detection using Deep Neural Network in Intel Architecture. Retrieved from http://www.digit.in
[9] https://d2l.ai/chapter_computer-vision/ssd.html#bounding-box-prediction-layer.ss



ISSN: 2233-7853 IJFGCN