International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 7 ǁ July 2022 ǁ PP. 85-90

Crowd Control and Monitoring using SSD

Abhijith Lal[1], Ajayakrishnan J[2], Amritha S[3], Josmi K Jose[4], Avani S[5]

[1]PG Student, Department of Computer Applications - Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala India

[2]PG Student, Department of Computer Applications - Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala India

[3]PG Student, Department of Computer Applications - Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala India

[4]PG Student, Department of Computer Applications - Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala India

[5]Assistant Professor, Department of Computer Applications - Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala India

ABSTRACT

Counting people in visual surveillance is hard and challenging problem. Automatic counting surveillance of individuals in public area is vital for safety control. Previously many techniques and methods are proposed.These methods/techniques aren't producing accurate and high performance for difficult situations. Now Foreground Extraction and Expectation Maximization (EM) based methods are proposed, which provides a far better accurate solution for counting people and locating a private . This work presents the security precaution of covid-19 for maintaining social distancing. Single shot detector algorithm(SSD) takes the live stream from camera and convolutional neural network(CNN) will identify the human and assign a private id and therefore the count it accordingly. In this work we have built a system that can detect and count the human entering a building and leaving a building. Using this data, we can count the number of people entered in a particular building, thus maintaining the Covid protocols for the limit of people in a particular functions and so on. This work has provided a 90% accurate output for the output.

Keywords: *Single shot detector(SSD) Algorithm, Deep Learning, Machine Learning, Convolutional neural network(CNN) Algorithm, Live stream*

Date of Submission: 24-06-2022 Date of acceptance: 05-07-2022

I. INTRODUCTION

People counting has a wide range of applications in the context of pervasive systems. These applications range from efficient allocation of resources in smart buildings to handling emergency situations. There exist several vision based algorithms for people counting. Each algorithm performs differently in terms of efficiency, flexibility and accuracy for different indoor scenarios. Hence, evaluating these algorithms with respect to different application scenarios, environment conditions and camera orientations will provide a better choice for actual deployment.

In this work, we are getting to build the Human Detection and Counting System through an uploaded video or Webcam . This is often an intermediate level deep learning work on computer vision, which can assist you to master the concepts and cause you to an expert within the field of knowledge Science.

II. LITERATURE SURVEY

A general example-based framework for detecting objects in static images by components is proposed in [1] . The technique is demonstrated by developing a system that locates people in cluttered scenes. Especially, the system detects the components of a person’s body in a picture, i.e., the head, the left and right arms, and therefore the legs, rather than the complete body by using four distinct example based detectors. The system then checks to make sure that the detected components are within the proper geometric configuration.

An image registrationalgorithm which recovers affine motion between a pair ofimages was presented in [2] . The algorithm proposed uses a log-polartransformation of the image pair to recover translational, rotational and scale misalignment.

Outlining and improvement of human tracking and detection for security applications were found in [3]. The hardware of the suggested work system including both the mechanical part and the embedded system is based on easily available materials, making it a cheaper option for low-budget products or applications. We are able to redefine it with n number of modifications like with 3D cameras also with face recognition or object recognition.

DCNN and Motion-based model for human detection in the infrared video feed were introduced in [4]. It shows the overall system architecture design which they used in their research. First, they reduced the noise effect for the overall system by applying the pre-processing on raw images. Raw images contained noises due to camera effect and background illumination variation. Noise filters and histogram equalization were used in the preprocessing stage.

III. METHODOLOGY

3.1 Single Shot Detector Algorithm

SSD is designed for object detection in real-time. Faster R-CNN uses a region proposal network to create boundary boxes and utilizes those boxes to classify objects. While it is considered the start-of-the-art in accuracy, the whole process runs at 7 frames per second. Far below what real-time processing needs. SSD speeds up the process by eliminating the need for the region proposal network. To recover the drop in accuracy, SSD applies a few improvements including multi-scale features and default boxes. These improvements allow SSD to match the Faster R-CNN's accuracy using lower resolution images, which further pushes the speed higher. According to the following comparison, it achieves the real-time processing speed and even beats the accuracy of the Faster R-CNN.

As said above the SSD model detects objects in a single pass, which means it saves a lot of time. But at the same time, the SSD model also seems to have amazing accuracy in its detection.In order to achieve high detection accuracy, the SSD model produces predictions at different scales from the feature maps of different scales and explicitly separates predictions by aspect ratio.

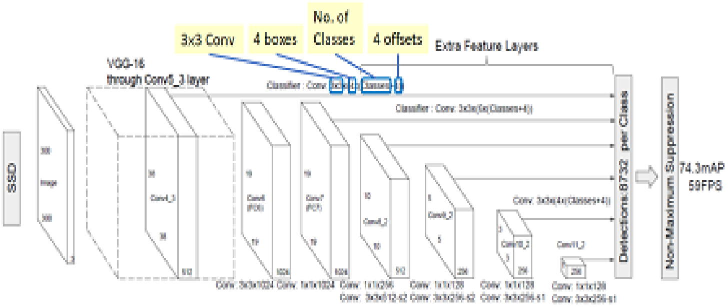


Fig 1 : Single Shot Detector

SSD has two components: a backbone model and SSD head. Backbone model usually is a pre-trained image classification network as a feature extractor. This is typically a network like ResNet trained on ImageNet from which the final fully connected classification layer has been removed. We are thus left with a deep neural network that is able to extract semantic meaning from the input image while preserving the spatial structure of the image albeit at a lower resolution. For ResNet34, the backbone results in a 256 7x7 feature maps for an input image. We will explain what feature and feature map are later on. The SSD head is just one or more convolutional layers added to this backbone and the outputs are interpreted as the bounding boxes and classes of objects in the spatial location of the final layers activation.

3.2 Convolution Neural Network

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

The steps used in CNN are:

- First, we take a picture, video, Live CCTVcameras input.
- Then we divide the image into various regions.
- We'll then consider each region as a separate image.
- Pass of these regions (images) to the CNN and classify them into various classes.

- Once we've divided each region into its corresponding class, we will combine of these regions to urge the first image with the detected objects.
- The matter with using this approach is that the objects within the image can have different aspect ratios and spatial locations. as an example, in some cases the thing could be covering most of the image, while in others the thing might only be covering a little percentage of the image. The shapes of the objects may additionally vary (happens tons in real-life usecases)

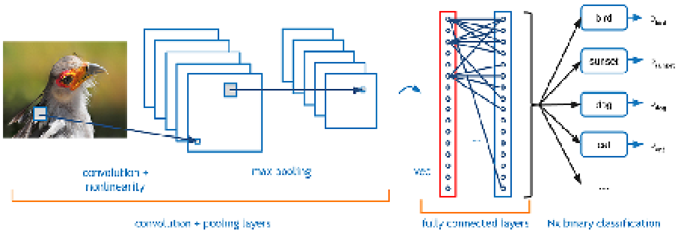
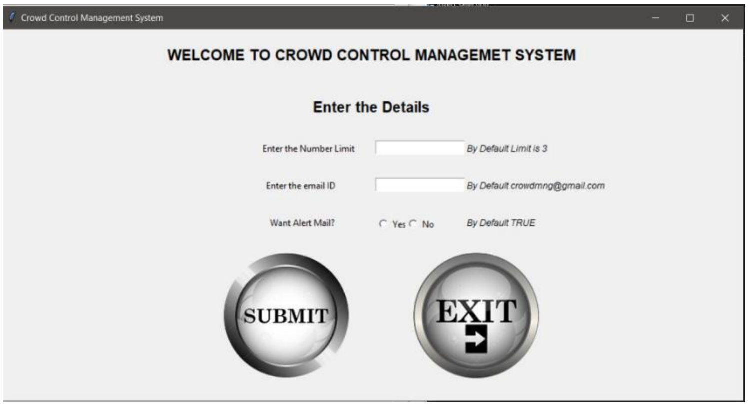


Fig 2: Convolution Neural Network

IV. EXPERIMENT AND RESULTS

To experiment this work, we start with GUI(Fig 3) and entered necessary inputs such as number of limit to set the human count limit in a particular building, email id to send email alert to the provided email id and an option to set alert email. If the option is yes, then it will give an alert sound as well as alert email, it the option is no, then it will provide only alert sound.


Crowd Control Management System

WELCOME TO CROWD CONTROL MANAGEMET SYSTEM

Enter the Details

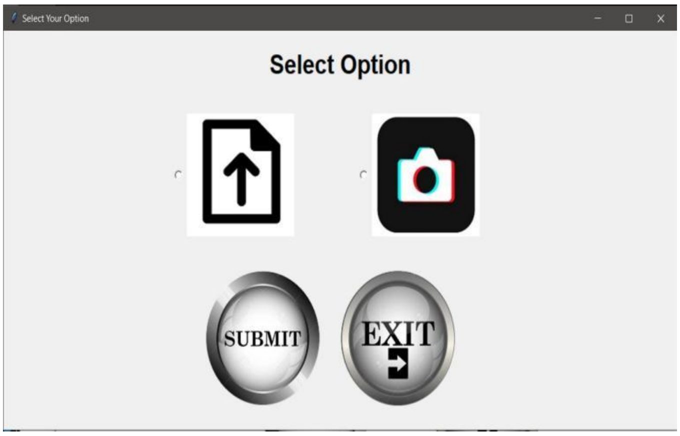
Enter the Number Limit *By Default Limit is 3*

Enter the email ID *By Default crowdmng@gmail.com*

Want Alert Mail? ○ Yes ○ No *By Default TRUE*

Fig 3: Welcome GUI

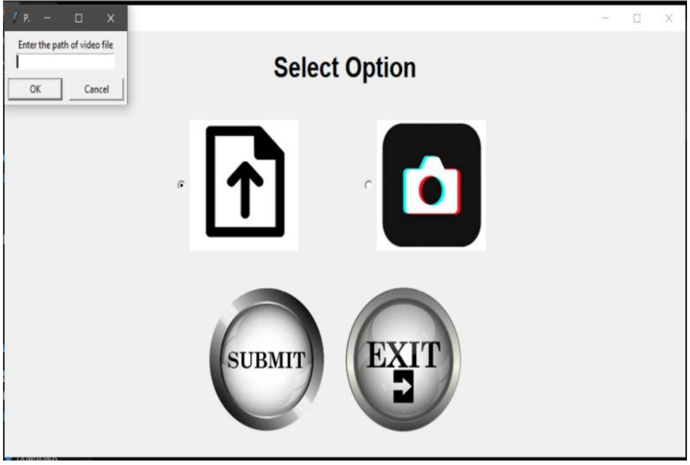
Once the submit button is pressed, it will direct you to a page where you will have option to select whether to upload a video or to open the web camera as shown in Fig 4. If you need to count the people count from a video, you can upload the video using upload button, else you can use web cam button.


Select Your Option

Select Option

Fig 4: Option screen

If you select upload, it will ask to enter the path of video as shown in the Fig:5



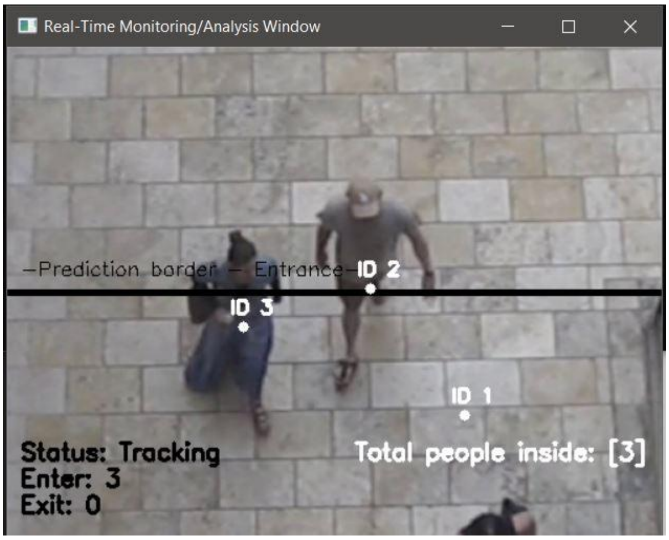
Enter the path of video file

Fig 5: Enter path

The video will be opened once entering the path of the video file and start tracking the video for human detection and counting as shown in the figure Fig: 6.

If you select the web camera option, then it will open your web camera and start tracking the human and count the humans as shown in the figure Fig:7.

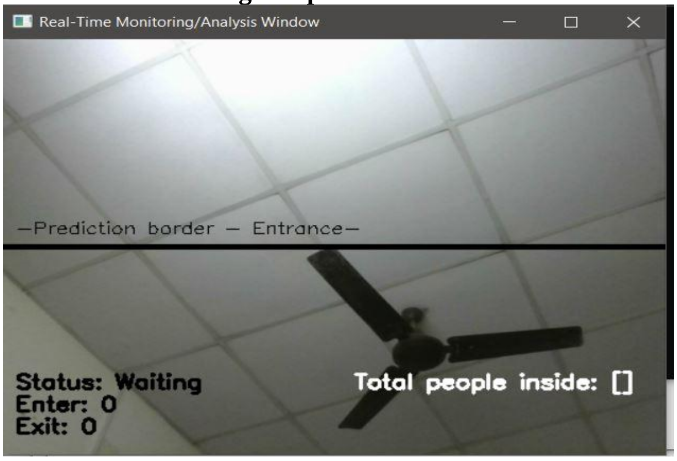
In both cases, we use a horizontal line to divide the screen and those who crosses the horizontal line is counted as one.



Real-Time Monitoring/Analysis Window
—Prediction border — Entrance—
ID 2
ID 3
ID 1

Status: Tracking
Enter: 3
Exit: 0
Total people inside: [3]

Fig 6: Uploaded video

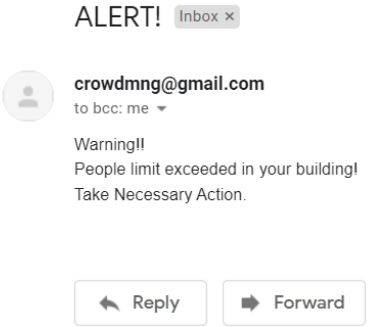


Real-Time Monitoring/Analysis Window
—Prediction border — Entrance—

Status: Waiting
Enter: 0
Exit: 0
Total people inside: []

Fig 7: Web Camera

If the limit exceeds the value entered in limit, then it will provide an alarm sound as well as sent an email with alert. Sample email attachment is shown in Fig:8



ALERT!

crowdmng@gmail.com
to bcc: me

Warning!!
People limit exceeded in your building!
Take Necessary Action.

Fig 8: Alert email

4.1 Test Report

Report was generated for this work is described below. It has tested all the test cases that is used for this work.

Table 1: Test Report for Crowd Control and Management System

Test Case #	Test Case Type	Description	Test Step	Expected Result	Status
1	Upload Video	An email has to be sent to the provided email address if the human count in the uploaded Video exceeds the limit	Upload a video from the path, provide an email and a human limit. Then Set up the Alert ON	When the human count exceeds the provided limit, an alarm sound will be generated and email is sent to the provided email address	Pass
2	Live Camera	An email has to be sent to the provided email address if the human count in the CCTV Camera/ Web Camera exceeds the limit	Open the CCTV or Web Camera.Provide an email and a human limit. Then Set up the Alert ON	When the human count exceeds the provided limit, an alarm sound will be generated and email is sent to the provided email address	Pass

V. CONCLUSION AND FUTURE SCOPE

Accurately detecting citizens during a visual closed circuit television is critical for a variety of applications such as abnormal event detection, human gait characterization, congestion analysis, person identification, gender classification, and fall detection for the elderly.

In this work, we have developed a system, that can detect and count the humans with 90% accuracy using SSD algorithm. The aim of this paper is to show that human detection and counting is most useful in this pandemic COVID-19 situation to maintain social distance in the public places by setting the people limit. As a result of which we can easily monitor the people limit in this COVID situation.

This System can be improved in the functionality to make it work from airport,railway stations or bus stands. As the COVID pandemic increases, this software is more needed in the upcoming days and the demand for this system will be increased. It can be improved for working from other places.

REFERENCES

[1]. Oren, M., Papageorgiou, C., Sinha, P., Osuma, E., Poggio,Publication year: 1997 “Pedestrian detection victimization wave templates”.

[2]. G. Wolberg and S. Zokai. Published year: 2000 “Robust Image Registration Using Log-Polar Transform”.

[3]. B. Karthikeyan M.E, Lakshmanan R, KabilanM.and Madeshwaran R. Published year: 2020 “Real-time detection and tracking of human based on image processing with laser pointing.”.

[4]. Heshan Fernando, Indika Perera and Chathura de Silva. Published year : 2019 “Real-time Human Detection and Tracking in Infrared Video Feed”.

[5]. E. P. Myint and M. M. Sein. Published year: 2021 “People Detecting and Counting System”.

[6]. A. R. Shahzad and A. Jalal. Published year: 2021 “A Smart Surveillance System for Pedestrian Tracking and Counting using Template Matching"