International Journal of Recent Advances in Multidisciplinary Topics
Volume 3, Issue 5, May 2022
https://www.ijramt.com | ISSN (Online): 2582-7839

Crowd Counting and Detection

Manorma Malik[1], Manu Sharma[2*], Navya Chopra[3]
[1,2,3]*Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India*

***Abstract*: Crowd counting is a valuable tool in today's culture. It has a wide range of applications, including human identification, population censuses, and security. Hu man tracking is a one-of-a kind image processing approach with a bright future. As a consequence of developments in deep learning, artificial intelligence, and other technologies, the number of crowds has increased substantially in recent years. The goal of this research is to create software that can monitor objects (people) as well as handle lists and counts. Using YOLOv5 (You Only Look Once) technology with Pytorch, the targeting system recognises objects and uses Deep Sort for tracking and counting. Furthermore, unlike the popular Yolo object recognition engine. which detects all items at once, this system recognizes just the objects required by the user, which aids in system speed.**

***Keywords*: Machine Learning, Deep Learning, Yolov5, DeepSort, Tracking.**

1. Introduction

One of the technologies that will be necessary in the future world is tracking. Multiple Object Tracking and single object tracking are the two primary categories of tracking. Object tracking is critical in the solution of several fundamental computer vision issues [1]. Detecting items in separate frames and blending them over several frames is required when tracking multiple objects in films. Many Computer Vision techniques are used to create object detection systems, and the technology is rapidly changing and opening up new possibilities. The image processing is completed by providing a labelled dataset that is trained and used as a model for the system, which can detect objects (humans) in several frames by comparing them to the objects provided within the model by mapping the model's same pattern inside the frame [2] YOLO (You Only Look Once), OpenCV, PyTorch, dataset, TKINTER, and GPU are some of the methodologies used to recognise, count, and track objects in Crowd Detection. To detect the objects, the proposed system employs the most recent Yolov5. For both training and detection, following object detection, the OpenCV module is used to input real-time or file format video input. is built using Pytorch Classifier in deep learning, and following object detection, the OpenCV module is used to input real-time or file format video input, as well as track and count the objects discovered inside the output, resulting in a highly efficient Crowd Detection system. Tkinter also makes group detection simple for user interaction, making it simple to choose a model that meets the user's needs, and only the real model object (person) is identified. It detects Humans within the crowd using Deep Learning Convolutional Neural Network models and so tracks it using the DeepSORT algorithm and provides the density of the group. the group Detection makes users take count of the identical form of humans and also won't detect them from the majority crowd helping users to avoid wasting time in searching. YoloV5 uses a PyTorch classifier. Yolo started with darknet technology and develop into Yolo 2, Yolo 3, and lastly Yolo v4 [3]. Yolov5 is a new app that was just released for easy tasks detection construction, resulting in improved object detection performance.

2. Related Work

A. Object Detection

The detection of objects is divided into three stages: categorization, detection, and segmentation [1]. Classification is the process of labelling a picture with a unique identifier in order to identify it during object detection. The next phase is detection, in which the item is detected in frames using the trained model. This delivers an honest view of the system providing the picture not ion. This stage recognises the articles with the model and delivers the object's position in this frame. The discovered item is then characterized and segmented for better comprehension in the following stage, segmentation (which usually uses a coordinate representation of a rectangular detection box).

B. YOLO

Joseph Redmon was the one who first proposed the YOLO algorithm. Alexey Bochkovskiy produced a paper on YOLO while he was first working on the YOLO algorithm and there was no major development from another author. After that, a succession of YOIO occurred, leading to YOLOv2, YOLOv3, and finally YOLOV4. On May 30, 2020, the Ultralytics LLC team launched YOLOVS in the wake of YOLOV4. YOLOV5 transitioned from the darknet to PyTorch, achieving 140 FPS in the Tesla P100, whereas YOLOv4 only achieved 60 FPS. 50 FPS. whose architecture is almost the same as that of YOLOv4. However, YOLOv5 does actual learning and object detection compared to YOLOv4.

C. DeepSort

DeepSort is a person-tracking machine learning algorithm that provides each individual a unique identifier. Sorting has the drawback of assigning a separate ID to a person who hides

*Corresponding author: manu83909@gmail.com

behind an object and then returns [5]. The procedure for DeepSort is as follows: Using YOLO, calculate the bounding tax (detection) To connect bounding boxes and tracks, use sorting (K alman filter) and ID (discriminative model). If the link cannot be formed, a new ID will be provided and the track will be re-added.

D. OpenCV

Opency is a free and open-source computer vision framework that includes image processing and camera access [7]. The GPU is currently included in the Opencv module. This is a key component of pytorch [7], and it helps to improve the efficiency of various algorithms, particularly in the field of image processing. Numpy, matplot, and utilis are among the libraries supported by Python's Opencv. Video image merging, navigation, and medical analysis are just a few of the applications for Opencv.

3. Proposed System

The system starts with a login screen, following which you may choose from live camera feeds, pictures, or stored video as input for object identification. After that, the data is processed into the frames for the detection process. We are detecting the objects(humans) by using Yolo V5 and trained Model in the detection process in the inputted data. Using the Yolo V5 after detecting objects(human), The Deep Sort algorithm provides the unique ID to the detected objects(humans) and count the accurate counting and density of the crowd. Then it is processed into output in inputted format by bounding boxes around the detected objects (humans) and their ID number.

A. Method Input

The programming library for computer vision and artificial intelligence is OpenCV (Open-Source Computer Vision Library). We can access the camera module and add video and image files in various formats. Real-time video and image frames from camera lenses are collected by OpenCV [7]. We assigned Tkinter window prompt before a collecting the real-time video and image frames, which represents or collects information from the user about the dataset or model that should be used for detection.

After the user provides an input option, the option is sent to the neural network module, and the camera is enabled using OpenCV to begin collecting video and image frames from the camera lens.

B. Tkinter

To create a tkinter app [4].

Import a module - tkinter

Create main window (container)

Add any number of widgets to the main window Apply Trigger event on widget

C. Integration

The five components listed above are utilized to create a single system. The system has a home page at the front. OpenCV is mostly used in a secure environment. Following the completion of the modules, it is converted into an executable file in the ".exe" file format, which can be used as a system output.

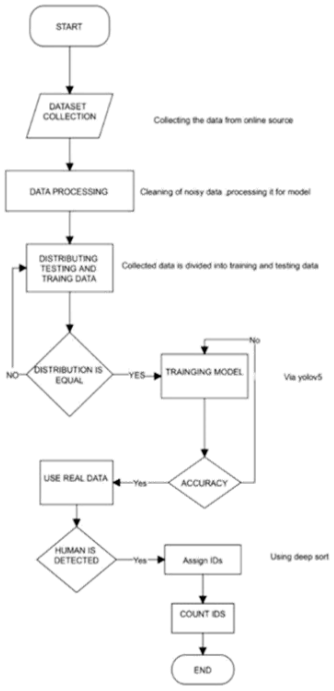

Fig. 1. Flowchart

4. Evaluation and Result of Performance

Three terms are used to assess YOLO's performance: mAP, Precision, and Recall [3]. mAP is a combined measure of recall and accuracy used to identify object accuracy. It is derived using a mean precision value for recall greater than 0 to 1 and an IOU that is the intersection of the two values between 0.5 and 0.95.

$$AP = (\sum Pr) = 1 / R \quad (1)$$

$$mAP = 1/N(\sum) \quad (2)$$

Precision is used to measure how accurately the objects are predicted. Precision can describe how good the model is at predicting the positive class.

$$Precision = (TP)/(TP+FP) \quad (3)$$

The Recall determines the detection quality of the objects or classes. Meaning the proportional recovery or sensitivity of the real positives are found correctly.

$$loU = Area\ of\ overlap/Area\ of\ union \quad (4)$$

YOLO accuracy can be found using mAP. Defines the interpolated AP used in the calculation. AP corresponds to average AP for loU from 0.5 to 0.95 in 0.05 step.

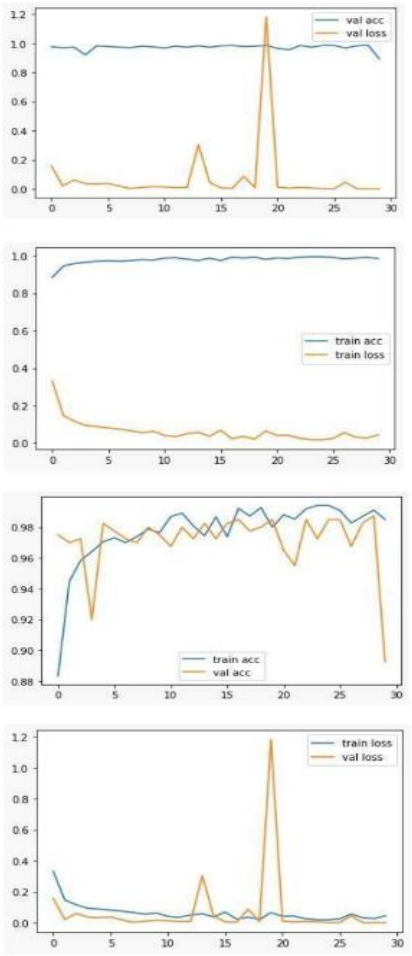

Fig. 2.

5. Conclusion

In this paper, we offer a crowd counting method. Yolov5-based technique for detecting trained participants as well as keeping track of and counting things in each frame. The crowd counting system contains a number of real-time features. Detecting certain items from a distance, for example. Monitoring a certain species of animal in a cluttered environment.

Object, or detecting a collection of feature classes, or counting a set of features, classes a specific item This system is extremely efficient in terms of Object detection on CPU GPUs. This system in place requires a local machine with a GPU, which may or may not be available, ideal for all systems, but especially with cloud-based sources like as GPU in Google Colab and also develop a functioning crowd-counting system capable of generating benefactors on a one-of-a-kind basis as a result, employing the yolov5. This system's algorithm allows it to keep track of a variety of things.

References

[1] L. Tan, X. Dong, Y. Ma and C. Yu, "A Multiple Object Tracking Algorithm Based on YOLO Detection," 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2018, pp. 1-5.

[2] H. Karunasekera, H. Wang and H. Zhang, "Multiple Object Tracking with Attention to Appearance, Structure, Motion and Size," in IEEE Access, vol. 7, pp. 104423-104434, 2019.

[3] C. G. Abhinu et al., “Multiple Object Tracking using Deep Learning with YOLO V5,” in International Journal of Engineering Research & Technology, vol. 9, no. 13, pp. 47-51, 2021.

[4] Tkinter, python, 2022, [Accessed on 29 Dec. 2021].

[5] DeepSORT: Deep Learning to Track Custom Objects in a Video, 2021. [Accessed on 29 Dec. 2021].

[6] Alzubaidi,L Zhang, Humaidi, A J et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J Big Data 8, 53(2021).

[7] Shiv Onkar Deepak Kumar, “Image Processing and Computer Vision with Python & OpenCV,” 2019.