IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 7.047 ⋇ Vol. 10, Issue 6, June 2022

DOI: 10.17148/IJIREEICE.2022.10625

Object Detection and Tracking using OpenCV

Nandini A[1], Preethi MP[2], Mrs. Kavitha Patil[3]

[1,2]UG Student, Department Of Information Science And Engineering, Atria Institute Of Technology, Bangalore, Karnataka, India.

[3]Assistant Professor in the Atria Institute of Technology's Department of Information Science and Engineering, Bangalore, Karnataka, India.

Abstract: A well-known application of computer vision and image processing, object detection focuses on finding objects or instances of those items in digital photos and videos. Face detection, character recognition, and vehicle calculation are just a few of the well-researched object detection applications. There are several uses for object detection, including retrieval and surveillance.

A wide, active, and challenging field of computer vision is real-time object detection. Image localization refers to the process of finding a single object in an image, while object detection refers to the process of finding several objects in an image. This recognises a class of semantic items in digital photos and movies.

INTRODUCTION

Recognizing and locating all recognised things in a scene is the goal of object detection. A long-standing technical aspiration of humanity has been to give machines intelligence and create increasingly independent and autonomous robots. If a robot cannot recognise and adjust to a changing environment, it cannot be too intelligent.

Real-time searching and recognition are exceedingly challenging tasks. This issue still hasn't had an efficient remedy identified for it. Despite extensive study in this field, the approaches that have so far been created are ineffective, call for lengthy training sessions, cannot be used in the real world, and cannot scale to a large number of classes.

LITERATURE SURVEY

The real-world traffic environment's illumination and weather influence, along with the onboard camera's motion blur, bumps, etc., make the bounding box susceptible to flickering and missing targets. The main tasks in every computer vision-based traffic monitoring system are the detection and precise localisation of the position of moving vehicles in a video series. A moving vehicle localization and bounding box estimation algorithm is put forth in this research. The adaptive background subtraction approach is used to separate the moving vehicle foreground from the static background, and the two-dimensional binary histogram projection profile (2D-BHPP) algorithm is used to estimate the bounding box of the moving vehicles. Prior to applying the 2D-BHPP algorithm, the foreground object refining is carried out via morphological closure operations. only the suggested approach[1]

Missed and erroneous detections are frequently caused by the bounding box flickering and missing targets. In order to complete the task of MOT on single-camera films, this study gives a thorough survey of works that use Deep Learning models. An extensive analysis of how Deep Learning was used in each of these stages is offered. Four key steps in MOT algorithms are highlighted. The three MOT Challenge datasets were used in a thorough experimental evaluation of the studies that were presented, which revealed many similarities among the best methods and suggested some potential future research topics.[2]

Recognizing the actual movement of related video degree in certain objects in specified video frames is the primary goal of motion-based object tracking. This gives a summary of the techniques currently in use for tracking moving objects, as well as information on the global rise in relevance of this subject. It falls short in terms of identifying strong patterns for better execution outcomes. The purpose of this study is to present an overview of the current approaches for the detection of moving objects as well as information about the increased global importance of this subject. In order to ensure that the target item is moving in a video that was captured by a single stationary camera, temporary information, mostly in the form of BG removal and temporal and frame differentiation, was seen as the primary methodology used.[3]

This paper provides a detailed review of deep learning-based object detection frameworks that address various sub-problems, such as occlusion, clutter, and low resolution, with varying degrees of modifications to R-CNN. Generic object identification pipelines, which serve as the foundation for other related tasks, are where the review begins. Following that, a quick discussion of three additional typical tasks follows, including pedestrian identification,

DOI: 10.17148/IJIREEICE.2022.10625

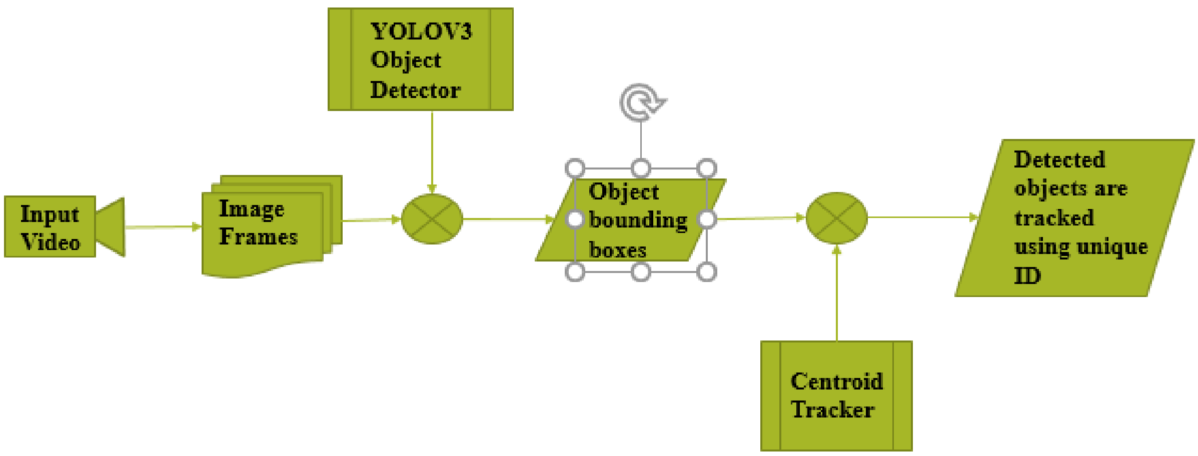
face detection, and prominent item detection. The advancements in neural networks and related learning systems, which offer priceless insights and directives for future growth, are also significant in the context of this review.[4]

This essay utilised To study motion of objects that is not visible to the naked eye, concepts like optical flow and histogram of magnitudes are used. Classification and localisation let the campus environment distinguish between normal and aberrant events, enabling the detection of both types of events. The main goal is to implement the straightforward process that will be used for the benchmark in the future. The task is divided into three main sections: zone division of the video frames, optical flow computation in each zone, and analysis and classification of the data as normal and abnormal occurrences using a logical threshold.[5]

Pretrained networks are used to extract features, while SVM is used to separate the outcomes of classification. Approach aids in directing ITS's course. The theoretical network is customised to automatically detect the level of congestion on the roads in order to facilitate the implementation of deep learning technologies in transportation applications. To connect the current cutting-edge deep learning algorithms with practical implementation, Traffic Net is presented. Three deep Traffic Nets based on residual learning are set up and pre-trained by the ILSVRC-2012 dataset after testing and training iterations. The traffic application is then given access to the network structure and weight. A dataset of traffic photos is created from live surveillance camera feeds in order to fine-tune the Traffic Net. However, the new input photos' detection accuracy is not as good.[6]

In this study, the use of a Tile convolution neural network and its recursive mode of operation aid in the discovery of objects useful for driver assistance system applications (DAS). To learn and adjust weights based on a variety of training data, the approach involves unsupervised training. To lower the number of valid detections, obstacle validation procedures are added. To produce reliable on-road vehicle and pedestrian identification, recognition, and tracking, the proposed system combines a unique deep learning approach with the utilisation of several sources of local patterns and depth information. The proposed system's first component, an effective adaptive U-V disparity algorithm, is based on robust obstacle detection to recognise impediments that are likely to be moving cars and pedestrians.[7]

ARCHITECTURE DIAGRAM



IMPLEMENTATION

The system starts with a login screen, then after choosing an option, takes live video or records video to feed into the object detection algorithm. After that, frames are created from the inputted data for detection. The Yolo V5 uses the Model or dataset during detection to find the object in the supplied data. The model is used to detect things, classify and represent the identified objects by placing bounding boxes around them and then processing the data into an output video format.

Classification, detection, and segmentation are the three phases of object detection. In order to identify the image during object detection, classification involves assigning it a specific id. The next phase is detection, where an object is found in frames using the trained model. This gives the system that provides the concept about the image a good perception. Using the model, this stage finds the locations of the items in the frame and detects them.

Joseph RedMon is the yolo algorithm's original author. When he first began building the yolo algorithm, another author, Alexey Bochkovskiy, released an article on it when it didn't make much progress. Subsequently, a succession of yolo appeared, leading to yolov2, yolov3, and finally yolov4. On May 30, 2020, the Ultralytics LLC team released YOLOV5

ISO 3297:2007 Certified ⵣ Impact Factor 7.047 ⵣ Vol. 10, Issue 6, June 2022

DOI: 10.17148/IJIREEICE.2022.10625

in the midst of yolov4. Yolov5 switched from the darknet to Pytorch, achieving 140 FPS in the Tesla P100 as opposed to Yolov4's 50 FPS. The advantages and architecture of Yolo V5 are virtually identical to those of Yolo V4. However, Yolov5 is more practical for training and object detection than Yolov4 in comparison.

RESULT

This approach used towards increase the object detection and tracking more simple and efficiently. YOLO'S speed and accuracy makes it a widely used algorithm for real time object detection. Using centroid tracker one or more moving objects are tracked simultaneously with unique id in given input video

1. Input

We may add video files in a variety of formats and access the camera module using OpenCV. The real-time video frame is gathered from camera lenses using OpenCV. A Tkinter window prompt that asks the user to indicate or provide information about which dataset or model should be used for detection was assigned before the real-time video frame was collected. The user option is given to the neural network module as soon as the input option is received, and OpenCV is then used to enable the camera so that it may begin gathering video frames from the camera lens.

YOLO(V5)

Each frame is given to the yolo detection algorithm with the model that the user selects after the input data from the camera or video is classified into frames. A new model can be developed for detection, or the model can be one that is predefined, such as the COCO dataset model. Once the item has been located, the detection is bound with boxes and labels and sent to the output portion, where the detected frames are gathered and then compressed into output format. The recognised frames are used for tracking, counting, and sorting using OpenCV prior to merging, as well as for better outcomes. Object tracking and sorting are both done with DeepSORT.

2. Data set

In order to create a custom model that may be utilised for detection, this field is used to create a custom dataset from raw photos. In order to do this, a dataset is first created by gathering the raw photos from multiple sources. The items must then be identified and labelled from the dataset photographs.

To annotate and identify the objects in this case, Python frameworks like "Labeling" are utilised. The dataset is divided into train and test images in proportions of 70% for train and 30% for test when annotation and labelling are complete, as this is the typical optimal percentage utilised for training. Once this is completed, the dataset can be sent to the Yolo training algorithm.

CONCLUSION

With the help of this thesis and based on experimental findings, we may more accurately detect objects and identify each one with its precise location in the image on the x and y axes. This research examines the efficacy of each method for item detection and identification and provides experimental findings for several strategies.

This approach used towards increase the object detection and tracking more simple and efficiently.

REFERENCES

[1] ."Fast approach for moving vehicle localization and bounding box estimation in highway traffic videos**"** Harikrishnan, P.M.; Thomas, A.; Gopi, V.P.; Palanisamy, P(2021)
[2] "Deep learning in video multi-object tracking**"** Ciaparrone, G.; Sánchez, F.L.; Tabik, S.; Troiano, L.; Tagliaferri, R.; Herrera, F(2020)
[3] "A Motion based Object Tracking System" Bhunesh Patel, Neel Ray, Priyanka Patel(**2020**)
[4] "Object Detection with Deep Learning" Zhao, Z.-Q.; Zheng, P.; Xu, S.-T.; Wu, X(2019)
[5] "Detecting Abnormal Events in University Areas" Zahraa Kain(2018).
[6] "Detection of unwanted traffic congestion based on existing surveillance system using in freeway via a CNN-architecture traffic net" P. Wang(2018)
[7] "Learning Framework for Robust Obstacle Detection, Recognition, and Tracking" V. D. Nguyen (2017)