International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 03 | Mar 2022 www.irjet.net p-ISSN: 2395-0072

VEHICLE DETECTION USING YOLO V3 FOR COUNTING THEVEHICLES AND TRAFFIC ANALYSIS

CH. Ranga Reddy[#1], Angshuman Roy[#2], Ms. G. Priyanka[#3], Dr. G. Victo Sudha George[#4]

[#1-#4]*Dept. of Computer Science & Engineering, Dr. MGR Educational and Research Institute, Chennai, India*

ABSTRACT: Joining of the advanced innovations in the observation of metropolitan versatility, for example, checking the traffic thickness will help in working on the amount of vehicles/plans to be accommodated the public compensation, office to be fused in decreasing the traffic, foundation to be given, for example, street extending, person on foot way, over span, underpass and so on, where traffic and transport is an issue. This can be executed in the city, at which it is perceived to be created as a savvy city. The proposed research work dissects the vehicle thickness utilizing python OpenCV and YOLOv3. Constant recordings are recorded in four ways from Sony HD IP cameras in an assigned region. Picture outlines from video succession are utilized to distinguish moving vehicles. The foundation extraction technique is applied for each casing which is utilized in ensuing investigation to recognize and count every one of the vehicles.

The masses are recognized for every vehicle which assists with following the vehicle moving. The focal point of every vehicle with mass gives the count of vehicle dependent on the paths considered. This work includes the vehicle progressively as well as groups the various vehicles utilizing profound learning strategy. YoloV3 (You just look once) object identification framework is utilized alongside a pre trained model called dark net to arrange the vehicle into various classes (transport, vehicle, cruiser and so on) This profound learning technique showed better grouping and recognition rate contrasted with masses and morphological strategy utilized for counting the vehicles. Arrangement is displayed for vehicle and further more individual grouping is considered to dissect the level of individuals and vehicles. The examination of level of vehicles is shown utilizing pie outline.

INTRODUCTION

In the present situation of quickly developing traffic condition, keeping up with the legitimate transportation framework is an extremely challenging position. It is because of the outstanding development of vehicles every day. Subsequently there is a need of robotized street transportation component absent a lot of human intercession. One of the fundamental components is to computerize the vehicle counting framework which offers clever vehicle observation framework. It helps in tracking down the present status of traffic and furthermore for overseeing it. This computerization helps in checking and assessing the continuous traffic stream in any area. Thus, it is one of the significant methods for streamlining the traffic lights. There are such countless techniques for assessing the traffic thickness, some of them are infrared or inductive circle indicators, radar and traffic cameras. Among these techniques, PC vision based strategy could be suitable since different strategies might have less execution and high upkeep. Be that as it may, PC vision based strategies additionally have downsides as a result of climate and lighting circumstances, HD video handling overhead and so forth, Intelligent transportation framework is one of the critical components for the improvement of savvy urban communities. This can be accomplished by examining the ongoing traffic thickness through video handling. This paper proposes a technique to find the traffic thickness continuously utilizing python-OpenCV library for video handling. Profound learning with pre- prepared model is additionally used to order the vehicles.

RELATED WORK :

Existing framework have grown continuous vehicle identification and counting from complex traffic situations utilizing low-rank deterioration with foundation subtraction(removal). This framework has shown normal execution and change location benchmark. Picture outlines from video arrangement are utilized to recognize moving vehicles. The foundation extraction technique is applied for each casing which is utilized in ensuing examination to recognize and count every one of the vehicles. The masses are distinguished for every vehicle which assists with following the vehicle moving. This profound learning technique showed better grouping and recognition rate contrasted with masses and morphological strategy utilized for counting the vehicles. Characterization is displayed for vehicle and furthermore individual arrangement is considered to break down the level of individuals and vehicles.

Yi-Qi Huang et al., has proposed in this paperIn the insightful rush hour gridlock framework, ongoing and precise discoveriesof vehicles in pictures and video information are vital and testing work. Particularly in circumstances with complex scenes, various models, and high thickness, it is hard to precisely find and arrange these vehicles during traffic streams. Subsequently, we propose a solitary stage profound neural organization YOLOv3-DL, which depends on the Tensorflow structure to work on this issue. The organization structure is streamlined by presenting the possibility of spatial pyramid pooling, then, at that point, the misfortune work is reclassified, and a weight regularization strategy is presented, for that, the continuous discoveries and measurements of traffic streams can be carried out successfully. The streamlining calculation we use is the DL-CAR informational collection for start to finish network preparing and tries different things with informational indexes under various situations and weathers.[1]

Muhammad Fachrie et al., has proposed inthis paper Deep Learning is a famous Machine Learning calculation that is broadly utilized in numerous spaces in current day to day existence. Its vigorous exhibition and prepared to-utilize structures and designsempowers many individuals to foster differentDeep Learning-based programming or frameworks to help human errands and exercises. Traffic observing is one region that uses Deep Learning for a considerable length of time. By utilizing cameras introduced in certain spots on the streets, many undertakings, for example, vehicle counting, vehicle recognizable proof, criminal traffic offense observing, vehicle speed checking, and so on can be figured it out. In this paper, we talk about a Deep Learning execution to make a vehicle counting framework without following the vehicles developments. To improve the framework execution and to decrease time in conveying Deep Learning engineering, thus pretrained model of YOLOv3 is utilized in this examination because of its great presentation and moderate computational time in object discovery. This exploration intends to make a basic vehicle including framework to assist human with ordering and counting the vehicles that go across the road. The counting depends on four kinds of vehicle, for example vehicle, cruiser, transport, and truck, while past researchcounts the vehicle only.[2]

Jun Liu and Rui Zhang et al., has proposed in this paper Vehicle identification is an essential errand for independent driving and requests high precision and constant speed. Taking into account that the current profound learning object identification model size is too huge to even think about being sent on the vehicle, this paper acquaints the lightweight organization with adjust the elementextraction layer of YOLOv3 and further develop the excess convolution structure, and the further developed Lightweight YOLO network lessens the quantity of organization boundaries to a quarter. Then, at that point, the tag is distinguished to work out the genuine vehicle width and the distance between the vehicles is assessed by the width. This paper proposes a recognition and going combination technique dependent on two diverse central length cameras to tackle the issue of troublesome location and low precision brought about by a little tag when the distance is far away. [3]

Adel Ammar et al., has proposed in this paper In this paper, we address the issue of vehicle location from ethereal pictures utilizing Convolutional Neural Networks (CNN). This issue presents extra difficulties when contrasted with vehicle (or any article) identification from ground pictures since components of vehicles from flying pictures are more hard to perceive. To explore this issue, we evaluate the exhibition of two cutting edge CNN calculations, to be specific Faster R- CNN, which is the most famous locale based calculation, and YOLOv3, which is known to be the quickest recognition calculation. We examine two datasets with various attributes to really look at the effect ofdifferent variables, like UAV's height, camera goal, and item size. The goal of this work is todirect a vigorous correlation between these two state of the art calculations. By utilizing an assortment of measurements, we show that none of the two calculations outflanks the other in all cases.Unmanned flying vehicles (UAVs) are these days a critical empowering innovation for an enormous number of utilizations like observation, following, catastrophe the executives, brilliant leaving, Intelligent Transport Systems, to give some examples. Because of their adaptability, UAVs offer novel abilities to gather visual information utilizing high- goal cameras from various areas, points, and altitudes..[4]

Haoxiang Liang et al., has proposed in this paper Intelligent vehicle location and including are turning out to be progressively significant in the field of parkway the board.

Notwithstanding, because of the various sizes of vehicles, their identification stays a test that straightforwardly influences the precision of vehicle counts. To resolve this issue, this paper proposes a dream based vehicle recognition and counting framework.. Contrasted and the current public datasets, the proposed dataset contains commented onsmall articles in the picture, which gives the total information establishment to vehicle discovery dependent on profound learning. In the proposed vehicle identification and counting framework, the parkway street surface in the picture is first extricated and separated into a distant region and a proximal region by a recently proposed division technique; the strategy is pivotal for further developing vehicle location. Then, at that point, the over two regions are set into the YOLOv3 organization to distinguish the sort and area of the vehicle. [5]

PROPOSED SYSTEM :

With out the utilization of the any equipment We propose YOLO V3 calculation to recognize and count the complete number of vehicles.This technique utilizes just go for it v3 and picture handling library to distinguish and count the quantity of vehicles.YOLO models give better outcomes acquired to pivoting objects, little objects.YOLOv3 (You Only Look Once, Version 3) is a continuous article recognition calculation that distinguishes explicit items in recordings, live feeds, or images.High pace of vehicle thickness and count can be gotten with the most extreme accuracy.YOLO v3 utilizes avariation of (Darknet is an open source neural organization system written in C. upholds CPU and GPU calculation), which initially has 53 layer network prepared on Imagenet. For the assignment of discovery, more layers are stacked onto it, giving us a layer completely convolutional basic design for YOLO v3.

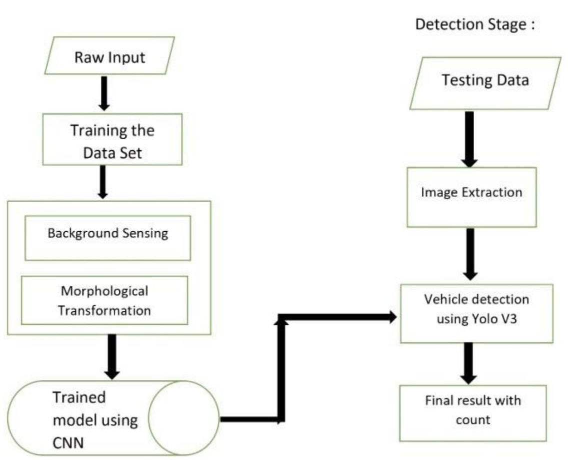

Fig 1

EXPERIMENTAL SETUP

Inputs and outputs:

A batch of images is given as input where each image has a shape (m, 608, 608, 3).

- A list of bounding boxes is outputted with the recognized classes. Every bounding box is represented with 6 numbers which are as follows pc, bx, by, bh, bw, c. If c is expanded into a vector of 80-dimensions, each bounding box will then be represented by 85 numbers.

Anchor Boxes:

- After exploring the traning data, the reasonable ratio of height/ width is chosen, that will represent the different classes.

- In the encoding m, nH, nW, anchors, classes, the dimensions of anchor boxes is the second to last dimension.

The YOLO architecture is: IMAGE (m, 608, 608, 3) (as input will be provided to the deep CNN)--> DEEP CNN (then it will be encoded)--> ENCODING (m, 19, 19, 5, 85).

Encoding:

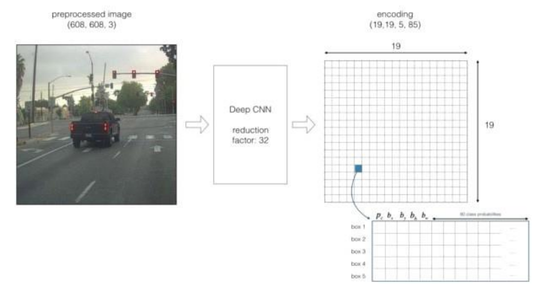

Fig 2

Let's take a look in great detail at what this encoding represents. A grid cell becomes responsible for detecting an object, when the centre/midpoint of an object falls into a grid cell. As we are making use of 5 anchor boxes, every 19 x 19 cells encodes information about 5 boxes. Height and width are the only defining properties of an Anchor box. In order to simplify things, we will flat line the last two last dimensions of the shape in the (19, 19, 5, 85) encoding. So the output of the Deep CNN came out to be (19, 19,425).

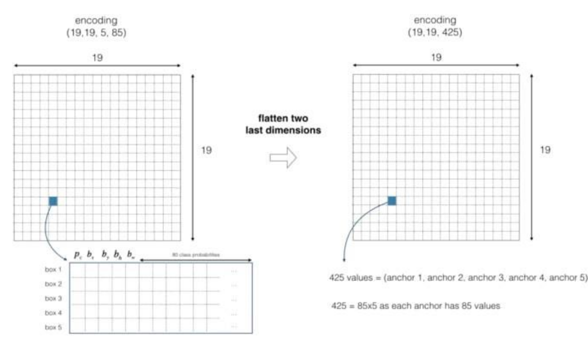

Fig 3

Class score:

Computing the element-wise productto extract a probability that the box contains a certain class will be done on each box (ofeach cell).

The class score is score_c_i = p_c * c_i — the probability that there are an object p_c times the probability that the object is a certain classc_i.

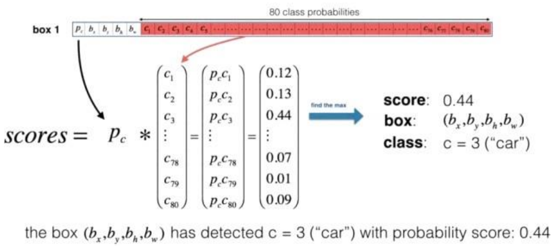

Fig 4

In Figure 4, let's consider box 1 (cell 1), the probability of an object existing in p_1 is =

0.60. So we can conclude that there's a 60% chance that an object exists in box 1 (cell 1).

The probability of an object being present in the class category 3 (a car) is c_3 = 0.73.

The score for box 1 and for category 3 is score $c_{1,3}$ = 0.60 * 0.73 = 0.44.

Consider we calculate the score for all 80 classes in box 1 and it turns out that the score for the car class (class 3) is the maximum. So, box 1 will be assigned to class 3 and a score of 0.44.

Visualizing classes:

Here's a way to visualize the prediction of YOLO algorithm in an image:

We will find the maximum of the probability scores for each of the 19 x 19 grid cells, i.e, we will take a maximum score across 80 classes, one maximum score for each of the 5 anchor boxes.

The algorithm will colour the grid cells according to what it considers to be the most likely object.

Doing this results in this picture:

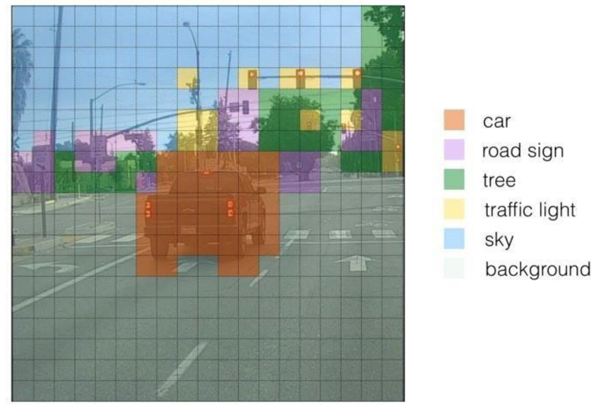

Fig 5

It should be noted that YOLO algorithm's core part is not this visualization itself formaking the predictions. It's a way of visualizing the intermediate result of the algorithm.

Visualizing bounding boxes:

Plotting bounding boxes around the output is another way to visualize the YOLO's output.

Doing this outcomes into a visualization like this:

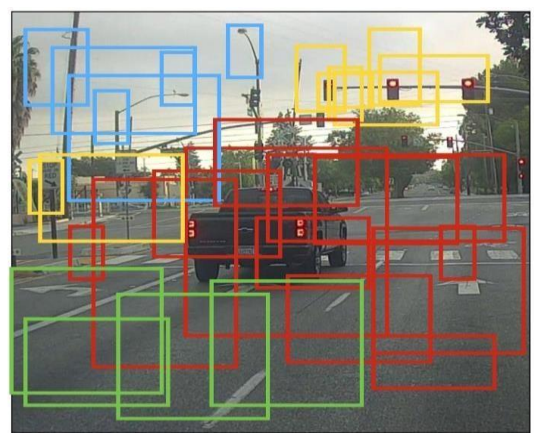

Fig 6

Non-Max suppression:

In the figure above, boxes are plotted around the objects for which the model had assigned a high probability, but this is still too many boxes. In order to create a better and accurate output, we will reduce the algorithm's output to a smaller number of detected objects.

To achieve that, we'll use non-max suppression. We'll carry out the following steps:

We will get rid of boxes with a lower scores. Meaning, the box is not very confident about detecting a class; either due to the low probability of any object or the low probability of this particular class.

We will select only one box when severalboxes will overlap with each other to detect the same object.

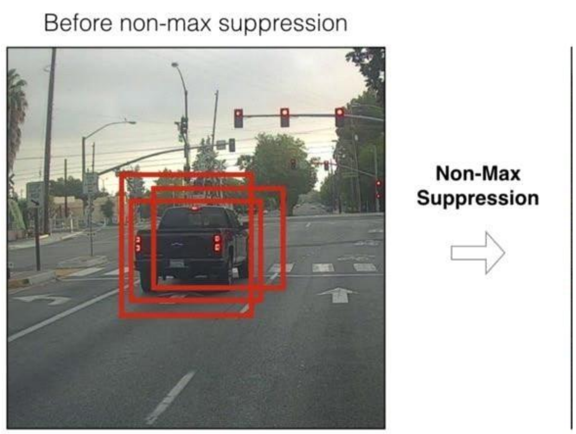

Fig 7

Filtering with a threshold onclass scores:

A filter by thresholding is applied. We will first apply a filter by thresholding. We well then get rid of any and all boxes for which theclass "score" is less than a chosen threshold.

The model outputs a total of 19 x 19 x 5 x 85 numbers, where each box is described by 85 numbers. It is often convenient to rearrange the (19, 19, 5, 85), or (19, 19, 425) dimensional tensor into the following variables:

box_confidence: It's a tensor of shape (19×19, 5, 1)(19×19, 5, 1) that contains p-c

— confidence probability, which means that there's a chance of some object being present

— inside each of the 5 boxes predicted ineach of the cells with dimensions 19x19.

— boxes: (19×19, 5, 4) shaped tensors that contains the midpoint and dimensions (bx, by, bh, bw) for each of the 5 boxes in each cell.

box_class_probs: tensor of shape (19×19, 5, 80) containing the "class probabilities" (c_1, c_2,...,c_{80}) for each of the 80 classes for each ofthe 5 boxes per cell.

Implementing yolo_filter_boxes():

- As seen in figure 4, we will compute the box scores by doing element-wise product(p*c).

- We will find the maximum box score, and the corresponding box score for each of the box.

- We will create a mask by using a threshold. As a heads-up: ([0.9, 0.3, 0.4, 0.5, 0.1] < 0.4) returns: The output of [False, True, False, False, True]. The mask should be "True" for the boxes we want to keep.

- Use Tensor Flow to apply the mask to box_class_scores, boxes and box_classes to filter out the boxes we don't want. It will leave us with subset of boxes we want to keep.

- The recognized information is utilized for counting the quantity of vehicles. The outcomes are contrasted and manual counting and it is shown. The outcome shows that the precision of the framework additionally relies upon the course wherein the point of the picture is given.

Algorithm	**Efficiency**
YOLOv3	**95**
CNN	**85**

CONCLUSION

There is probability of enormous number of constant utilizations of vehicle following and counting as a result of expanded traffic. In this paper, two techniques are analyzed and inferred that yolov3 is better for vehicle count and characterization. An investigation is given better perception utilizing pie-diagram showing the level of every class. This investigation helps for better traffic and group the board in complex rush hour gridlock situations (without the paths). This paper shows an intricate situation with examination of every arrangement for normal Indian streets with better precision from yolov3 model. It tends to be additionally reached out for auto vehicle discovery, speed estimation and course of development for every vehicle. The precision can be improved by cons idering the all conceivable vehicle classification.

REFERENCES

1. Yi-Qi Huang, Jia-Chun Zheng, Shi- Dan Sun , Cheng-Fu Yang and Jing Liu, Enhanced YOLOv3 Algorithm and Its Application in Traffic Flow Detections", Appl. Sci. 2020, 10, 3079,doi:10.3390/app1009307.

2. Muhammad Fachrie," A Simple Vehicle Counting System Using Deep Learning with YOLOv3 Model", DOI: 10.13140/RG.2.2.15026.56001, 2020.

3. Jun Liu and Rui Zhang, "Vehicle Detection and

Ranging Using Two Different Focal Length Cameras", Hindawi, Journal of Sensors, Volume 2020.

4. Adel Ammar1, Anis Koubaa1, Mohanned Ahmed1, AbdulrahmanSaad, "Ethereal Images Processing for Car Detection utilizing Convolutional Neural Networks: Comparison between Faster R- CNN and YoloV3", IEEE, 2020.

5. Huansheng Song, Haoxiang Liang, Huaiyu Li, Zhe Dai and Xu Yun, "Vision- based vehicle identification and including framework utilizing profound learning inparkway scenes", Springer openAccess 2019.

6. DonatoImpedovo, FabrizioBalducci , Vincenzo Dentamaro and Giuseppe Pirlo, "Vehicular Traffic Congestion Classification by Visual Features and Deep Learning Approaches: A Comparison", Sensors 2019, 19, 5213;doi:10.3390/s19235213.

7. Fukai Zhang, Ce Li and Feng Yang, "Vehicle Detection in Urban Traffic Surveillance Images Based on Convolutional Neural Networks with Feature Concatenation", Sensors 2019, 19, 594; doi:10.3390/s19030594.

8. Dinh Viet Sang, Duong Viet Hung, "YOLOv3-VD: A meager organization for vehicle identification utilizing variational dropout", SoICT 2019, December 4–6, 2019, Hanoi - Ha Long Bay, Viet Nam .

9. Lecheng Ouyang, Huali Wang, "Vehicle target identification in complex scenes dependent on YOLOv3 Algorithm", AMIMA 2019.

10. Manoharan, Samuel. "A further developed wellbeing calculation for man- made consciousness empowered processors in self driving vehicles." Journal of Artificial Intelligence 1, no. 02 (2019): 95-104.