International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

Object Detection Using YOLO Models

Muskan Choudhary[1], Sadanand Singh[2], Abhishek Kumar[3], Vinay Kasana[4], Nidhi Sharma[5]

[1,2,3,4] *Student, Computer Engineering Dept, Delhi Technical Campus, Greater Noida, Uttar Pradesh, India*
[5] *Professor, Computer Science Dept, Delhi Technical Campus, Greater Noida, Uttar Pradesh, India*

Abstract - *The world is currently shifting forward with rapid technological development, innovation and studies. This has helped to enhance human being's lives that will have a broader perspective. The present day technology in diverse drones and sensors has helped in numerous methods to get the item to meet the desires.*

Drones are broadly used ultra-modern world for plenty purposes. This includes taking pictures, live data consisting of live surveillance structures utilized by police or infantrymen to protect and secure regions under their control. It is consequently used as a surveillance device, this paper focusses on the improvement of the chance detection to ensure a faster more accurate and smaller human intervention model and subsequently proposes a model algorithm to use hazard reporting without the want for police intervention at the scene which makes the system absolutely independent. This will make casualty reporting less difficult and less complicated for the police and the person who is a part of the accident as there may be no need to call the police and there is no necessity to be physically present at the scene of the accident.

In this paper we suggest the usage of You Only Look Once (YOLO) algorithm for human like monitoring of roads from a bird's eye view.

Key Words: **YOLO; Object Detection; Computer Vision; Deep Learning**

1. INTRODUCTION

Road Traffic accidents have become very common nowadays. As majority of people are buying cars and other automobiles, the incidences of road accidents are simply growing each day. moreover, the roads have become narrower, and the towns have turn out to be greatly populated.

A total of 2,403 cases of road accidents happened on Expressways that caused injuries to 1,997 persons and deaths of 1,389 persons. The most number of causalities in road accidents were reported on the National Highways accounting for 34.4% (53,213 out of 1,54,732) followed by State Highways (25.6%) (39,624 deaths). Altogether 60,506 persons died due to road accidents on the other roads during 2019.

A total number of 25 patrolling vehicles are deployed on Agra Lucknow expressway which is a stretch of 302 km. In many cases the first responders are owners and employee of roadside eateries which works fine if an accident occurs nearby the same. Since most of the areas are isolated, this results in the lag of response time which being a crucial first step to the save the life of the people injured.

This scenario gives the rise to the need to monitor the roads 24/7 which can be achieved in an efficient manner using UAV's.

Currently, classical object detection strategies predicated on region proposals comprise of region-based Convolutional Neural Networks (R-CNNs), spatial Pyramid Pooling Networks (SPP-net), Fast R-CNNs, Faster R-CNNs, and Region-based Fully Convolutional Networks (R-FPN) . However, these approaches were futile in achieve concurrent speed because of the expensive running methods and incompetency of region propositions.

You only Look once (YOLO) is the most in demand object detection software program used in severa intelligent vision applications because of its simple use and high item detection accuracy. Additionally, in latest years, diverse clever vision structures based on high performance and overall performance inbuilt structures are being developed.

Although, the YOLO still requires high end hardware for booming real time detection. In this paper, we first discuss real-time object detection property of the YOLO models, in AI based systems with embedded structures with resource constraints.

Specifically, we tend to entail the issues associated with excessive precision and convenience of YOLO and provide real time object detection service by means of minimizing overall providers delay, that remains a limitation of the pure YOLO.

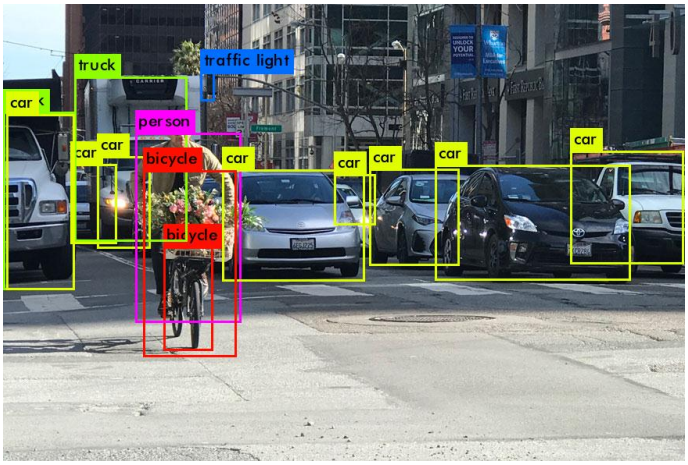

Figure 1 : Object Detection using Yolo algorithm

Thus, arise the need to enhance the pre-existing models of the YOLO algorithm. So, we will be comparing the fps and mAP of the current existing models of the same

2. LITERATURE SURVEY

Various research has been done on object detection from a vertical view. This view presents many challenges of its own few being smaller objects present in the background, viewpoint changes.

The number of computer vision (CV) tasks like object detection and image segmentation have gained extreme acceptance within the past few years. Object detection (OD) is difficult and helpful for determining the various visual objects of a particular class (such as cars, pedestrians, animals, terrains, etc.).

Object identification in this aspect has been a subject of interest for computer vision analysis for drone-based applications and autonomous navigation.

Though the best results can be seen by using two stage networks (YOLO + R-CNN) this results in loss of speed of detecting objects which is not favorable in the real-world scenarios since drone navigation and prediction should be quick to give accurate time bound result. *C. Liu et al.* [1] Image processing using YOLO network to enhance the detection of traffic signs.

The problem area which was highlighted is not good enough images being captured the problem of underexposure, blur, rotation persists. The work is based on realistic views of the real-world scenarios. The model is trained on the dataset of actual images to make the model more authentic for real world image detection. The YOLO neural network is used to analyze the object which are the traffic signs, and the analyses was based on Darknet-53 network structure. Furthermore, the analysis is done on certain factors namely Blur of the image, Flip or rotation, noise and cropping by using these to give the average precision of each factor. The comparison is done between the average precision that the test set gives versus the average precision of the real-world images captured by the model. The outcome shows a better object detection accuracy.

Traffic monitoring is not possible when done in a static way so to evolve this method of static watch *W. Fang et al.* [2] emphasis on object detection using YOLO, R-CNN, and DPM. The YOLOv3 version of YOLO has been discussed along with the advantages, drawbacks, and improvements of the YOLO algorithm. The characteristics of YOLO algorithm are that it is lightning fast and quick, global image specification is easy, it's easier to generalize and represent images etc.

The network design has been done as a convolutional neural network and the dataset used is PASCAL VOC. The network has been designed to maximize the extent of fast object detection. The data is trained and the measure adequately the detector identifies the locations and categories of objects throughout navigation is predicted with the help of logistics regression, further the object class is predicted, and the inference is generated for the image.

The limitations stated in this research comprise of the difference in results given by the huge variation in the size of the bounding box. Inaccurate restriction to a particular place is the major setback that has been observed due to the difference observed.

This work further narrates the future scope of the YOLO and YOLOv3 algorithm considering the COCO dataset which concludes the flexibility and accuracy achieved using the YOLOv3 over the formerly used algorithms when compared in areas of images detection and classification. When talking about real time surveillance *J. Tao et al.* [3] talks about the use of computer vision for object detection along with the deep learning modules particularly convolutional neural network and YOLO. The motive of this research is to identify and locate object in a traffic route, and to be used further for the purpose of surveillance of the traffic. The comparison between the traditional machine learning algorithms and deep learning algorithms for the prime purpose of object detection has been done and implementation shows the best outcome in the YOLO algorithm.

The need for real time surveillance and accurate results in different scenarios has led this research to happen. The recent related work done in this domain pointed to the use of convolutional neural network as the far best approach in getting the expected results. The factors of the designing scheme like network design, combining OYOLO and R-FCN and pre-processing the images are elaborated on. The experimentation includes the KITTI dataset and further explain the training process and the outcome achieved.

Adarsh et al. [4] explains the need of enhancements in the detection speed and accuracy has been a prime importance

when the monitoring traffic and providing accurate results for any causality using different object detection methods, which are HOG, RCNN, Fast RCNN, Faster RCNN, YOLO v1, YOLOv2, YOLOv3, SSD, etc.

The implementation and analysis give varied results for a particular characteristic like speed, accuracy, matching strategy, IOU threshold, training dataset, pace of learning, etc.

Using YOLOv3 tiny increases the speed of object detection in addition to better accuracy, on static object and on video containing dynamic pictures.

Image recognition is another factor that facilitates smooth working of the traffic so for that *Ratre et al.* [5] explains the YOLOv3 version of YOLO which has been discussed along with the advantages, drawbacks, and improvements of the YOLO algorithm. The characteristics of YOLO algorithm are that it is lightning fast and quick, global image specification is easy, it's easier to generalize and represent images etc.

The network design has been done as a convolutional neural network and the dataset used is PASCAL VOC. The network has been designed to maximize the extent of fast object detection. The data is trained and the measure how well the detector identifies the locations and classes of objects during navigation is predicted with the help of logistics regression, further the object class is predicted, and the inference is generated for the image.

The limitations stated in this research comprise of the difference in results given by the huge variation in the size of the bounding box. Inaccurate restriction to a particular place is the major setback that has been observed due to the difference observed.

This work further narrates the future scope of the YOLO and YOLOv3 algorithm considering the COCO dataset which concludes the flexibility and accuracy achieved using the YOLOv3 over the formerly used algorithms when compared in areas of images detection and classification.

But the YOLOv3 model conflicts with the FF-YOLO model based on accuracy, *L.Yitong et al.* [6] speaks on using the machine learning YOLO model for object detection in complex scenes by feature fusion. The work in this paper talks about the backdrop of YOLO V3 model in object detection and thus supersede it with FF-YOLO for the aim of faster and more accurate object detection in complex scenarios

For the needed improvement a four-scale detection layer in incorporated to the already existing three scale prediction mechanism for more precise input for an upgraded output.

Pascal VOC2007 and MS COCO data set are used in this paper for comparison of mAP on different targets. The comparison has been done between YOLOV3, YOLOV4, and FF YOLO models and the comparison has been done based on the number of targets detected and on those parameters the detection accuracy has been calculated which showed that FF YOLO gave results better than the rest in fuzzy or complex images with overlapped bounding boxes.

Xianbao et al. [7] improvised the YOLO v3 model to further refine the process of image detection.

Now we have the superseded version of YOLO algorithm which if FF YOLO by YOLO v3 owing to its better accuracy in detecting images but that itself does not help in monitoring the dynamic movements. *A.Sarda et al.* [8] takes a go on detecting objects in autonomous driving using YOLO and computer vision which eliminates the chance of mis happenings that might happen if human intervention exists. The model of YOLO used is YOLO V4 and the dataset used is OIDV4.

The detection rate of different algorithms is compared like the CNN, RNN, SVM, KNN which have been replaced by YOLO because of its high accuracy and faster results. The model which is being trained needs two prerequisites which are the most favorable coordinates of the bounding box and the object class. The work in this paper talks about using the YOLO neural network architecture in two of its models namely YOLOV3 and YOLOV4 along with mentioning about the upper hand that YOLOV4 model has over other models for data increase using synthetic data and other such techniques.

The data has been tested on 3200 images and trained on 8000 images. Furthermore, the module can be enhanced by training and testing on more data for precise results.

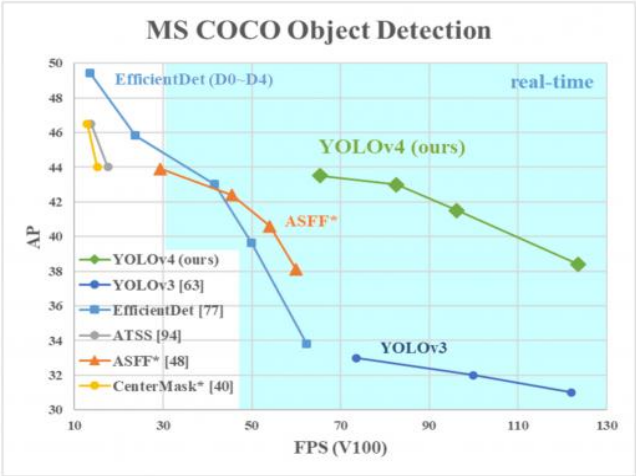

Figure 2 : Comparison of different object detection models

Figure 3 : Object Detection and precision

3. REQUIREMENT ANALYSIS

3.1 Deep Learning

Deep Learning is a subfield of machine learning and a pivotal part of artificial intelligence (AI). It is a neural network which attempt to simulate the behavior of human brain empowering it to learn from a large data set. It is an optimized neural network which shoes better result with an enhanced accuracy rate.

3.2 Artificial Intelligence

Artificial intelligence (AI) refers to the simulation of human intelligence in machines. The perfect characteristic of artificial intelligence is its ability to rationalize and take actions that have the simplest likelihood of achieving a selected goal

Artificial intelligence (AI) is the ability of a computer, or a droid controlled by a computer to try to do tasks that are done by humans as those need human intelligence and discernment.

3.3 YOLO V4

YOLO, as stated, stands for *You Only Look Once*, it is an object detection system in actual period that acknowledges various objects in an exceedingly single enclosure. Moreover, it identifies objects sooner and more exact than various recognition systems. YOLO is a futuristic recognizer that has a quicker FPS and is more precise than available detectors. The detector will be trained and used on a standard GPU that allows widespread adoption. New options in YOLOv4 improve accuracy of the classifier and detector and may be used for other research projects.

3.4 Python

Python is an interpreted, object-oriented, high-level programming language with dynamic linguistics developed by Guido van Rossum. It had been originally released in 1991.

It is a multiparadigm, all-purpose, interpreted, high-level programming language. Python permits programmers to use totally different programming designs to form easy or complicated programs, get faster results and write code nearly as if speaking in a very human language.

3.5 Open CV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and currently it plays a significant role in data processing that is extremely vital in today's systems. By using it, one can method pictures and videos to spot objects, faces, or perhaps handwriting of an individual. Once integrated with various libraries like NumPy, python is capable of processing the OpenCV array structure for analysis. To spot image pattern and its various features we have a tendency to use vector area and perform mathematical operations on these features.

4. CONCLUSION

Object detection for the use case we are trying to solve requires prediction at a real-time speed otherwise it is no better than manually monitoring drones. YOLO (V4) in particular gives a detection speed of 32 fps on devices with low GPS power and goes to almost 150 fps on graphic heavy devices. The drones these days have a maximum capacity of 4gb or graphic cards and with the introduction of powerful yet light chips (like Apple's M1 chip) it is only going to rise. So that way Yolo is the best algorithm at this time for detecting objects for a real time usage as object detections in drone

But there are still a lot of issues in a real time use case because most of the times the vehicle or object may be far away and would hamper the performance and accuracy but with time and variable speed of the drones can be able to solve this problem as well.

REFERENCES

[1] C. Liu, Y. Tao, J. Liang, K. Li and Y. Chen, "Object Detection Based on YOLO Network," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), 2018, pp. 799-803, doi: 10.1109/ITOEC.2018.8740604.

[2] W. Fang, L. Wang and P. Ren, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments," in IEEE Access,vol.8,pp.1935-1944,2020,doi: 10.1109/ACCESS.2019.2961959.

[3] J. Tao, H. Wang, X. Zhang, X. Li and H. Yang, "An object detection system based on YOLO in traffic scene," 2017 6th International Conference on Computer Science and Network Technology (ICCSNT), 2017, pp. 315-319, doi: 10.1109/ICCSNT.2017.8343709.

[4] Adarsh, P., Rathi, P. and Kumar, M., 2020, March. YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 687-694). IEEE.

[5] Sharma, A., Singh, A., Shetty, C. and Ratre, S., 2020, June. YOLO (You Only Look Once) Technology and Its' Impact in Field of Object Detection. In *Proceedings of the International Conference on Recent Advances in Computational Techniques (IC-RACT)*.

[6] C. Baoyuan, L. Yitong and S. Kun, "Research on Object Detection Method Based on FF-YOLO for Complex Scenes," in IEEE Access, vol. 9, pp. 127950-127960, 2021, doi: 10.1109/ACCESS.2021.3108398.

[7] Xianbao, C., Guihua, Q., Yu, J. *et al.* An improved small object detection method based on Yolo V3. *Pattern Anal Applic* **24,** 1347–1355 (2021).

[8] A. Sarda, S. Dixit and A. Bhan, "Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 1370-1374, doi: 10.1109/ICICV50876.2021.9388577.

[9]towardsdatascience.com/yolo-you-only-look-once 17f9280a47b0

[10]towardsdatascience.com/yolo-v4-or-yolo-v5-or-pp-yolo dad8e40f7109