ISSN No:-2456-2165

Analytical Study on Object Detection using Yolo Algorithm

Dawn Wilson[1], Dr. Manusankar C.[2], Dr. Prathibha P H[3]
PG Scholar[1], Assistant Professor and Head[2], Assistant Professor[3]
PG Department of Computer Science Sree Sankara Vidyapeetom college
Valayanchirangara, Perumbavoor Ernakulam, Kerala

Abstract:- Object detection is a technique that allows detecting and locating objects in videos and images. Object detection is widely used to count objects in a scene, track their precise locations and accurately label the objects. It seeks to answer what is the object? and Where is it? . Object detection adopts various approaches such as fast R-CNN, Retina-Net, Single Shot MultiBox Detector (SSD) and YOLO. Among these, YOLO is the most powerful algorithm for object detection and as well as suited for real-time scenarios. It is popular because of its accuracy and speed. YOLO uses Neural networks to provide object detection.

Keywords:- *YOLO, Neural Networks, CNN, Object detection.*

I. INTRODUCTION

Object detection is one of the interesting computer vision techniques. One of the major application of object detection is self-driving cars where systems use computer vision, technologies for the multidimensional representation of roads and LADAR. YOLO stands for 'You Only Look Once'. This is an algorithm that detects various objects in a picture (in real-time). Object detection in YOLO is a regression problem and use the class probabilities of the detected images.

YOLO algorithm uses convolutional neural networks (CNN) to detect objects in real-time. The algorithm requires only a single forward propagation through a neural network to detect objects. It is why it is named as 'You Only Look Once'. On a single run of the algorithm, the prediction of entire image is done. CNN is mainly used to predict bounding boxes and each class probabilities. Subsequent versions of YOLO include YOLO V1, YOLO V2, and YOLO V3 up to YOLO V5.

II. CONVOLUTIONAL NEURAL NETWORK

A Convolutional Neural Network (ConvNet/CNN) is a deep learning algorithm that can assign importance (learnable weights and biases) to various aspects/objects in an image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much less complicated than in other classification algorithms.

In fig 2.1 Convolutional Neural Networks are a different type of neural network. The layers are arranged in three dimensions: width, height, and depth. The neurons in one layer don't connect to all the neurons in the next layer, but only to a small region of it. The final output will be a vector of probability scores, each representing the probability of the observation occurring at a particular depth.

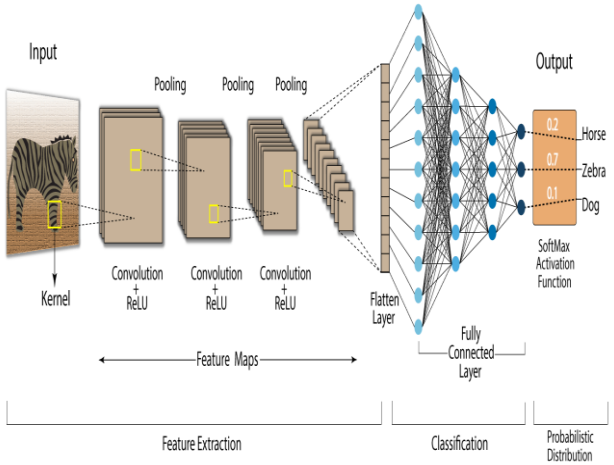

Fig. 2.1 Conventional Neural Network

III. WORKING OF YOLO

YOLO includes three techniques. They are Residual Blocks, Bounding Box Regression and Intersection Over Union[IOU].

A. Residual blocks

This is the first step. Here, the image fig:-3.1 is divided into some grids. Each grid has a S * S dimension.

Fig. 1: Grid Dimensions

ISSN No:-2456-2165

A bounding box is a box like outline to highlight objects. It consists of the following attributes:

- Width (bw)
- Height (bh)
- Class (example: car, person, etc.)- This is represented by the letter c.
- Here fig:-3.2 Bounding box center (bx,by)

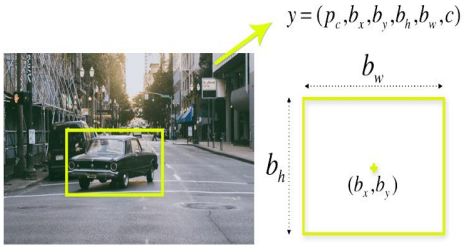

Fig. 2: Bounding Box

B. Intersection over union (IOU)

Intersection over union (IOU) describes how boxes overlap. In fig. 2 it provide an output box that surrounds the objects perfectly.

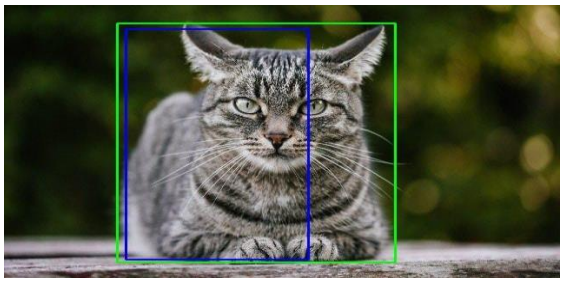

Fig. 3: IOU Box overlap

There are two bounding boxes in the above image, blue and green. The blue box is the predicted box while the green box is the real box. YOLO ensures that the two bounding boxes are equal.

IV. COMBINATION OF THE THREE TECHNIQUES

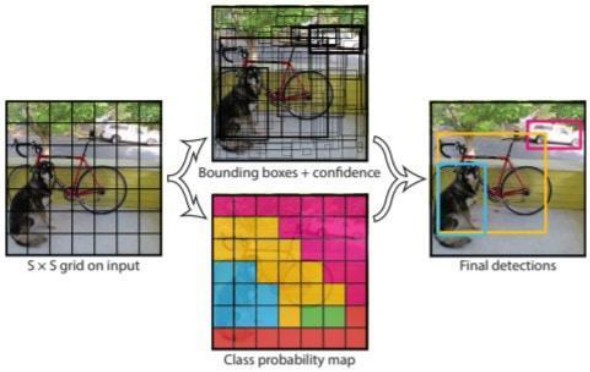

Fig. 4: single convolutional neural network

For example 4, There are at least three types of objects: a car, a dog, and a bicycle. All of the predictions are made simultaneously using a single convolutional neural network.

The intersection over union ensures that the predicted bounding boxes match the real boxes of the objects. This process eliminates unnecessary object bounding boxes that don't match the characteristics of the objects. The final detection will be made by fitting objects perfectly into predefined bounding boxes.

For example, The car is enclosed by the pink bounding box while the bicycle is enclosed by the yellow bounding box. The blue box has been used to highlight the dog.

V. LITERATURE SURVEY

The paper [1] The YOLO algorithm is used to detect and classify objects in images. It was trained on the classes from the COCO dataset. YOLO uses a single convolutional network to predict the likely boundaries of objects in an image Fig:-5.1. The bounding boxes are weighted according to probabilities, and the model uses the final weights to determine when to detect them. The bounding box represents four dimensions, including the center of the bounding box, width, height and 4. The value of 'c' refers to an object class.Flow chart of YOLO algorithm:

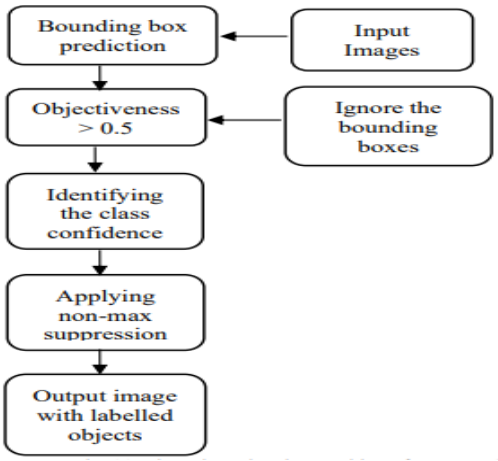

Fig. 5: Detect and classify objects

Applications of YOLO include Vehicle detection, Crowd Detection, Optical Character Detection and Image Fire Detection.

The paper [2] proposes a unified object detection algorithm using YOLO. They combined separate components of object detection into a single neural network. Fast YOLO is the fastest object detector currently available, and it is also one of the most advanced real-time object detectors out there. YOLO also generalizes well to new domains, making it ideal for applications that rely on fast, robust object detection. An experimental comparison with different object detection systems has shown that YOLO can be generalized to new domains better than other detectors.

The paper [3] this new network model suggests a modified version of the YOLOv1 network model that better predicts user behavior. The YOLOv1 network is improved by optimizing its 'e loss function. The inception model structure is added. A spatial pyramid pooling layer is used to pool data. The proposed model is effective at extracting features from images, outperforming other models in object detection. It is

trained using an end-to-end method, and has shown effective results on a challenging Pascal VOC dataset.

The paper [4] "YOLOv3: An Incremental Improvement" by Joseph Redmon and Ali Farhadi discusses how YOLOv3 is an incremental improvement over earlier versions of the algorithm. In this paper, the authors present some updates to the YOLO algorithm. They made a number of small design changes to improve it. We also trained a new network that's really great.It's a little bigger than last time, but more accurate.

The paper [5] "Comparative Research of YOLO Architecture Models in Book Detection" by Maria Kalinina and Pavel Nikolaev discusses the different YOLO architecture models and their effectiveness in detecting books. This white paper describes the creation of a deep convolutional neural network based on the YOLO architecture for real-time book detection.

The Paper [7] "Real-Time Objects" by Priya Kumari, Sonali Mitra, Suparna Biswas, Sunipa Roy, Sayan Roy Chowdhury, Antara Ghosal, Palashree Dhar and Anurima Majumdar. The main goal of this paper is to find the location of an object in a given picture accurately and mark the object with the appropriate category. We used a real-time object detection algorithm, YOLO, to train our machine learning model.YOLO is a powerful neural network that can detect objects in real time and with the help of the COCO dataset. This technique can detect an object in real time with 90% accuracy.

The paper [8] CAP-YOLO: Attention-based YOLO channel reduction for real-time Intelligent Monitoring. In this paper, CAP-YOLO (Channel Attention based Pruning YOLO) and AEPSM (adaptive image enhancement parameter selection module) are proposed to help achieve real-time intelligent analysis for coal mine surveillance videos.

VI. CONCLUSION AND FUTURE WORK

Object detection is a Computer Vision and Image Processing technique that deals with detecting instances of various classes of objects from the captured image or video. As compared to other object detection algorithms and approaches, YOLO is the fast and accurate detection algorithm. It follows a regression approach, that is, rather than selecting a field of interest for an image, they estimate groups and bounding boxes for the whole picture in one run of the algorithm [1].

Improvements are still making on YOLO. To obtain better results, the loss function of YOLO can be modified. Pooling layer can be added. The future and scope of YOLO object detection is vast

REFERENCES

[1.] Dr. S.V. Viraktamath, Rachita Byahatti, Madhuri Yavagal "Object Detection and Classification using YOLOv3" Dept. of Electronics and Communication Engineering SDM College Of Engineering and Technology Dharwad, India. International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV10IS020078 ,Vol. 10 Issue 02, February-2021

[2.] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi "You Only Look Once: Unified, Real-Time Object Detection", University of Washington , Allen Institute for AI , Facebook AI Research.

[3.] Tanvir Ahmad , Yinglong Ma , Muhammad Yahya, Belal Ahmad, Shah Nazir , and Amin ul Haq. Hindawi Scientific Programming Volume 2020, Article ID 8403262, 10 pages https://doi.org/10.1155/2020/8403262.

[4.] Joseph Redmon, Ali Farhadi, University of Washington, "YOLOv3: An Incremental Improvement"

[5.] Maria Kalinina, Pavel Nikolaev "Comparative Research of YOLO Architecture Models in Book Detection"

[6.] Priya Kumari1, Sonali Mitra, Suparna Biswas, Sunipa Roy, Sayan Roy, Chaudhuri, Antara Ghosal, Palasri Dhar, Anurima Majumder, "YOLO Algorithm Based Real-Time Object Detection", Volume 8 Issue 1

[7.] Zhi Xu , Jingzhao Li , Yifan Meng and Xiaoming Zhang "CAP-YOLO: Channel Attention Based Pruning YOLO for Coal Mine Real-Time Intelligent Monitoring"

[8.] https://towardsdatascience.com/yolo-you-only-look-once-3dbdbb608ec4

[9.] https://appsilon.com/object-detection-yolo-algorithm/