International Journal of Innovative Research in Engineering
Volume 3, Issue 4 (July-August 2022), PP: 67-72.
www.theijire.com
ISSN No: 2582-8746

Object Recognition from Video Using Yolo Algorithm

Dr.L.Malathi[1], Pooja.R[2], Sasipriya.B[3], Sushma.P[4], Susmitha.R[5]

[1]Professor, Computer Science and Engineering, Vivekanandha College of Engineering for Women, Tiruchengode, Namakkal, Tamilnadu, India.
[2,3,4,5]Student, Computer Science and Engineering, Vivekanandha College Of Engineering For Women, Tiruchengode, Namakkal, Tamilnadu, India.

How to cite this paper: Dr.L.Malathi[1] Pooja.R[2] Sasipriya.B[3] Sushma.P[4] Susmitha.R[5], "Object Recognition From Video Using Yolo Algorithm", IJIRE-V3I04-67-72.

Abstract: *Us humans can detect and recognise objects very quickly by using the senses of sight and touch. But when it comes to computers or machines, it is a very challenging one. Object detection is a computer vision technique for locating objects in images or videos. In other words, object detection is a supervised learning problem, which means that train our models on labelled examples. Each frame in the training dataset must be accompanied with a file that includes the boundaries and classes of the objects. It is useful in many fields like automatic car, military, etc. There are numerous calculations for identifying the items. The popular exciting algorithm of object detection are CNN, RCNN. But the main disadvantages are the accuracy rate is very low, when it comes to real time applications.To overcome the traditional CNN algorithm, additional features were added to the algorithm and it was named as YOLO. YOLO algorithm employs Convolutional Neural Network(CNN) to detect objects in real time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects. This means the prediction in the entire frame is done in a single algorithm run. YOLO algorithm accuracy rate and time complexity areefficient.In YOLO algorithm, YOLOv5 is a recent version. While comparing to other version, YOLOv5 is an efficient in application like object identification, fault detection, etc. This paper uses the YOLOv5 for recognizing images from video. It first split the videos into frames, that frames can be considered as images. Then it split the frames (i.e., images) by using the CNN (Convolutional Neural Network). By using the training dataset, the machine will analyse the objects in the frames. Then the output is displayed with the prediction of objects with highaccuracy.*

Key Word:*YOLO – You Look Only Once, object detection.*

I. INTRODUCTION

Human beings are one of the best creatures in the world. They have different types of senses and emotions. Emotions make humans understand each other. Likewise, senses make humans understand things. Sometimes, senses help them to save their life by responding to that situation. For example, if someone touches the fire which is glowing in the candle, he immediately takes his hand away from the fire. So, senses play a major role in human's life. There are five types of senses i.e., sight, touch, hear, smell and taste. Each sense helps in a different way. The sense of sight and touch are used to predict the object for the humans. Like humans, computers also are able to predict objects by using their computer vision. There are such countless applications in PC vision. Object detection is a subset of computer vision[2]. Object detection is used to detect the location, presence and types ofobjects.

II.MATERIAL AND METHODS

Existing System

RCNN might have been one of the principal huge and effective utilization of convolutional brain organizations to the issue of article limitation, recognition, and division.

It is a moderately basic and direct utilization of CNNs to the issue of article restriction and acknowledgment. A disadvantage of the methodology is that it is slow, requiring a CNN- put together element extraction pass with respect to every one of the competitor locales produced by the district proposition calculation.

The R-CNN models might be by and large more precise, yet the YOLO group of models are quick, a lot quicker than R-CNN, accomplishing object location progressively.

Proposed System

The YOLOv5 calculation includes a solitary brain network prepared start to finish that snaps a picture as info and predicts jumping boxes and class marks for each bouncing box straightforwardly. The method offers higher prescient exactness (for example low limitation mistakes), despite the fact that works at 45 edges each second and up to 155 casings each second for a speed-improved rendition of themodel.

The model works by initial parting the contribution to a network of casings, where each edge is liable for anticipating a jumping box on the off chance that the focal point of a bouncing box falls inside the casing. Each casing predicts a jumping box including the x, y coordinate and the width and level and the certainty. A class forecast is likewise founded on each casing.

System Implementation

This project divides into 4 modules which are,

a) Dataset collection and pre-processing.
b) Preparation of YOLO weights and training of thenetwork.
c) Prediction
d) Testing and PerformanceAnalysis.

a) Dataset Collection andPre-processing

- Image dataset is collected for the objects to bedetected.
- Around 100 to 350 images are needed per object. Images are downloaded from Google images and open images dataset which consists of 9 million images and it contains 600 classes ofobjects.

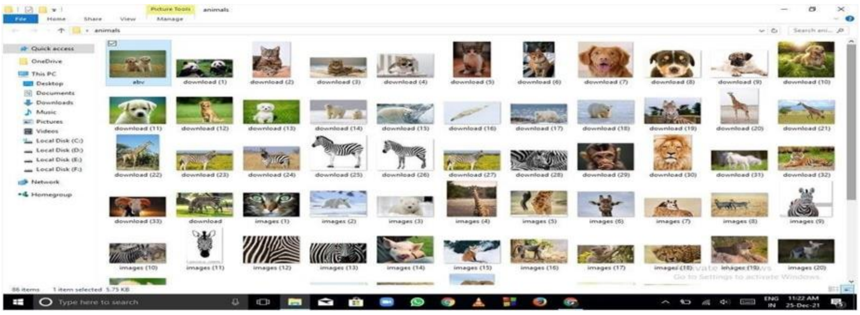


Fig 1.Image dataset.

Image dataset is prepared for every object. For example consider dog 120 images are collected for this specific object.

- They need to specify where the custom object is located on the specific image for training themodel.
- For this operation we will be using makesense.ai which is a graphical image annotation tool and label object bounding boxes inimages.

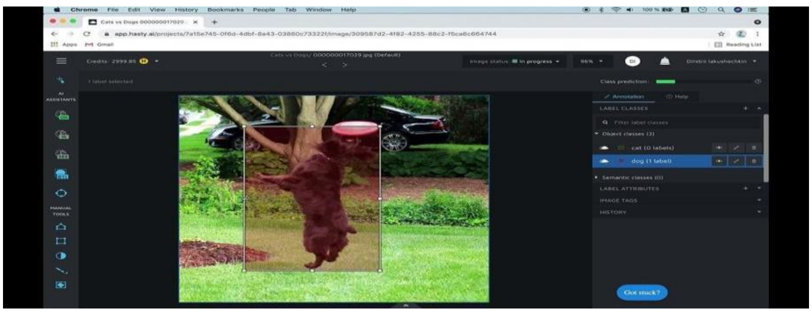


Fig 2.Sample image for dataset preparation.

Images are annotated in order to specify where the custom object is located on the image.

b) Preparation of YOLO weights and training of thenetwork.

- Set up Google drive and upload the collected dataset into thedrive.
- Type the required coding to train the model in GoogleColab.

- Connect the dataset with this file and train the model for few hours and download the weightsfile.

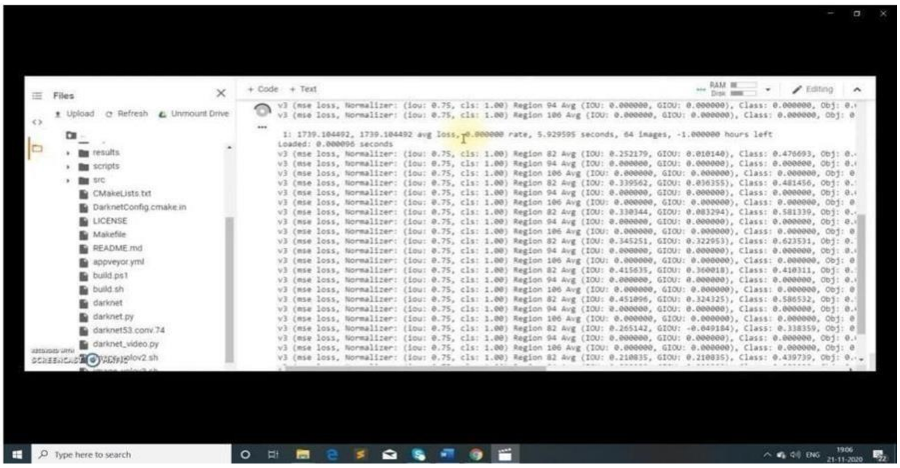


Fig 3.Training model

This is the screen displaying the training of the model in the Google colaboratory.

C) Prediction:

Labelled data:

After the completion of trained data, the machine will label the train data and store it in separate file. The below fig 4. shows the example of labelled dataset.

Fig 4.Predicted with Labelled dataset

The machine also predicts the confidence score for each and every trained dataset.
The below fig 5. shows the prediction of confidence score of the trained dataset.

Fig 5.Prediction of confidence score of the trained dataset

d) Testing and PerformanceAnalysis:

- Test the model by running the sourcecode.
- Check whether the object is correctly detected or not, along with the bounding box and classname.

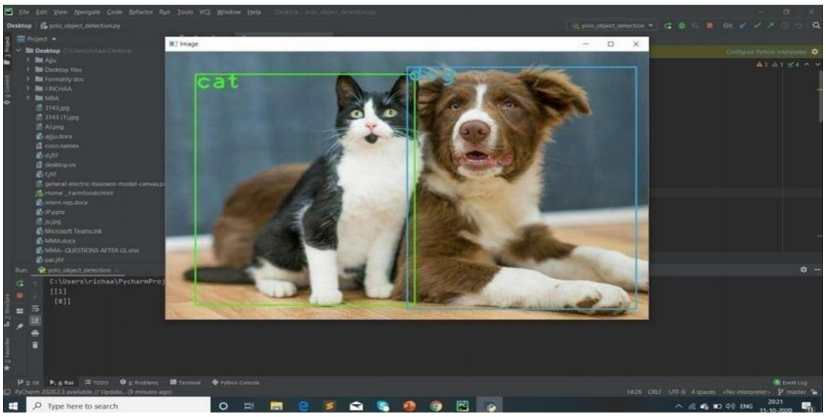


Fig 6.Tested data.

This is the expected output of the image containing the bounding boxes and class.

III.RESULT

Input Design

Input is a raw data that is processed to produce an output. A video which consists of one or more objects are given as input. After the completion of trained dataset, the machine will predict more than one object in the input video and saved in the particular file. The input can also be given by using the web camera, so the machine will predict the object in the real world entity.

Fig 7.Screenshot of Input video

Output Design

After the prediction of objects in input data, the machine will save the output in the particular folder. The below fig 8. shows the output design.

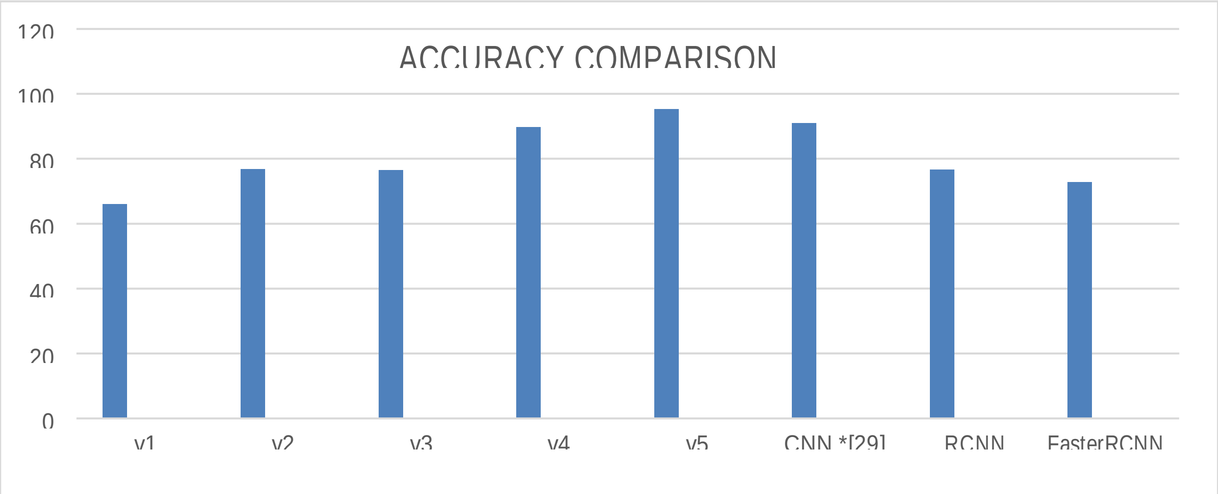
Fig 8. Screenshot of Output video

ALGORITHM	mAP	SPEED(ms)
CNN	49.87	2271
Faster RCNN	58.78	122
RCNN	68.98	79.000
YOLO V1	72.21	44.7
YOLO V2	74.67	45.1
YOLO V3	81.64	59.1
YOLO V4	83.16	59.6
YOLO V5	88.39	**25.2**

Table 1. Performance comparison of YOLO algorithm with other algorithm.

When comparing to other object detection algorithm, YOLOv5 takes a minimum time to run. So, it is easy to detect the objects for larger datasets.

Accuracy Comparison In Graph



The accuracy rate of YOLOv5 algorithm is 95.6%. But when comparing to other algorithm, YOLOV1 algorithm is YOLOv5 has high accuracy rate in real time applications.

The accuracy of YOLO V1 is 71%-66%, accuracy of YOLO V2 is 76.8%, accuracy of YOLOv3 is 76.5%-93.3%, accuracy of YOLOv4 is 89.7%, accuracy of CNN is 91%. When compared to other algorithms YOLO V5 has highest accuracy of95.3%.

IV.CONCLUSION

Due to its powerful learning ability and advantages in dealing with real world, machine leaning based object detection has become hotspot in recent years. This project include object detection methods, feature extraction, object classification. There are many traditional algorithms to detect objects. But the main disadvantage is that it is difficult to apply in real world applications. To overcome the drawbacks of the traditional algorithms, YOLOv5 was used.

In YOLOv5, it is easy to implement it in real world. When compared to other traditional algorithms, YOLOv5 has high accuracy and speed. It also takes minimum time to detect objects. It predicts the objects by using the trained datasets. The more we train the datasets, accuracy rate will be high. It also predicts the objects with the help of bounding boxes. While predicting the objects, it also displays the accuracy rate of objects. Our proposed system can be extended for CCTV security in surveillance.

References

[1] Object Detection through Modified YOLO Neural Network,Tanvir Ahmad, Yinglong Ma, Muhammad Yahya,Belal Ahmad,Shah Nazir,and Amin ul Haq, 2020.

[2] Visual Object Detection and Tracking using YOLO and SORT,Akansha Bathija M.Tech Student, Dept of Computer Engineering K J Somaiya College of Engineering Mumbai, Maharashtra, India,2019.

[3] A Practice for Object Detection Using YOLO Algorithm, Dr. Suwarna Gothane, Associate Professor, Department of Computer

Science and Engineering, CMR Technical Campus, Hyderabad, Telangana, India,April 2021,In DOI:10.32628/CSEIT217249

[4] https://medium.com/@venkatakrishna.jonnalagadda/object-detection-yolo-v1-v2-v3-c3d5eca2312a,2019.

[5] https://towardsdatascience.com/yolo-v4-or-yolo-v5-or-pp-yolo-dad8e40f7109,Chamidu Supeshala,2020.

[6] https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b,Ayoosh Kathuria, 2018

[7] https://ai-pool.com/a/s/yolov3-and-yolov4-in-object-detection,2021.

[8] https://towardsai.net/p/computer-vision/yolo-v5%E2%80%8A-%E2%80%8Aexplained- and-demystified,MihirRajput,2020.

[9] Real Time Video Analytics For Object Detecting And Face Identification Using Deep Learning(2019). Shrikant jagannath patro, Nisha V M.

[10] Survey Of Video Based Small Target Detection (2021), Ying Liu,LuyaoGeng, Weidong Zhang, and Yanchao Gong.53

[11]A Review Of Detection And Tracking Of Object From Image And Video Sequences (2017).Mukeshtiwari,Dr.Rakeshsinghai.

[12] Face Recognition And Retrieval In Video (2020). CaifengShan.

[13] Deep Learning For Object Detection And Scene Perception In Self-Driving Cars:Survey, Challenges, And Open Issues (2021). Abhishek Gupta, Alagananpalagan, Ling Guan, Ahmed Shaharyarkhwaja.

[14].Moving Object Detection For Video Surveillance(2015).K. Kalirajan And M.Sudha.

[15] .A Review on Video-Based Human Activity Recognition, by Shian-Ruke, Jang-HeeYoo, HoangLeUyenThuc, Yong-Jin Lee, Jeeng-Neng Hwang, KyoungHo Choi.

[16]A Moving Object Recognition Using Video Analytics Ambeshwar Kumardr. Rajesh, Assistant Professor Department Of Computer Science & Engineering Dayanandasagar University, Bangalore, India.

[17] A Review On Real Time Object Tracking In Video Sequences[2015]. Shraddhakothiya, ,Kinjal mistree.

[18] Vehicle Counting For Traffic Management System Using YOLO And Correlation Filter (2018).

[19] Refining Object Detection In Video With Recurrent Neural Networks (2016).

[20] Social Distancing Analyzer Using Computer Vision And Deep Learning, G V Shalini (2021).

[21] Investigations Of Object Detection In Images/Videos Using Various Deep Learning Techniques And Embedded Platforms—A Comprehensive Review(2020) 54 Chinthakindi Balaram Murthy, Mohammad Farukh Hashmi , Neeraj Dhanraj Bokde And Zong WooGeem.

[22] Object Detection And Tracking In Video Image(2014).Rajkamal KishorGupta.

[23] Object Detection With Deep Learning: A Review(2019).Zhong-Qiu Zhao, Peng Zheng, Shou-Tao Xu, And XindongWu.

[24] Fast Yolo: A Fast You Only Look Once System For Real-Time Embedded Object Detection In Vedio. Mohammad Javad Shafiee, Brendan Chywl, Francis Li, Alexander Wong.

[25] Real Time Object Detection And Trackingsystem For Video Surveillance System. Sudan Jha, Changho Seo, Eunmol Yang, Gyanendra PrasadJoshi.

[26] Human Detection In Thermal Videos Using Yolo. Marina Ivasic-Kos, Mate Kristo, MiranPobar.

[27] A Real-Time Object Detection Algorithm For Video. Shengyu Lu, Beizhan Wang, Hongji Wang, Lihao Chen, Ma Linjian, Xiaoyan Zhang.

[28] YOLO Based Detection And Classification Of Objects In Video Records. Arka Prava Jana; Abhiraj Biswas;Mohana.