International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 05 | May 2019 www.irjet.net p-ISSN: 2395-0072

# Object Detection using Machine Learning Technique

**Praahas Amin[1], Anushree B. S.[2], Bhavana B. Shetty[3], Kavya K.[4], Likitha Shetty[5]**

*[1]Assistant Professor, Dept. of Information Science Engineering, Sahyadri College of Engineering and Management, Karnataka, India*
*[2,3,4,5]Student, Dept. of Information Science Engineering, Sahyadri College of Engineering and Management, Karnataka, India*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** *The main aim of this project is to build a system that detects objects from the image or a stream of images given to the system in the form of previously recorded video or the real time input from the camera. Bounding boxes will be drawn around the objects that are being detected by the system. The system will also classify the object to the classes the object belongs. Python Programming and a Machine Learning Technique named YOLO (You Only Look Once) algorithm using Convolutional Neural Network is used for the Object Detection.*

***Key Words***: **Machine Learning, Object Detection, YOLO, Bounding Box, Convolutional Neural Network**

## 1. INTRODUCTION

This chapter describes the concepts used in the project Object Detection using Machine Learning elaborately.

The project comes under the domain Machine Learning which is the part of Artificial Neural Network. Machine Learning concepts makes the system learn on its own from the experiences it gains, without the interference of the external factors.

The YOLO (You Only Look Once) algorithm using Convolutional Neural Network is used for the detection purpose. It is a Deep Neural Network concept from Artificial Neural Network. Artificial Neural Network is inspired by the biological concept of Nervous System where the neurons are the nodes that form the network. Similarly, in Artificial Neural Network perceptrons act like the nodes in the network. Artificial Neural Network has three layers that are, Input Layer, Hidden Layer and the output Layer. Deep Learning is the part of the Artificial Neural Network that has multiple Hidden Layer that can be used for the Feature Extraction and Classification purposes.

Convolutional Neural Network (CNN) is the part of Deep Learning that is used in analysis of visual imagery. It has four different kinds of layers, they are, Convolutional Layer, Pooling Layer, Activation Layer and Fully Connected Layer. Convolution Layer uses filter and strides to obtain the Feature Maps. These Feature Maps are the matrix that is obtained after the Convolution Layer. It can be simplified using ReLU (Rectified Linear Unit) that maps negative values to 0. The resulted Feature Map is reduced by sending it into the Pooling Layer where it is reduced to the smaller sized matrix. This is how the features are extracted. At the end of the convolutional neural network is the Fully Connected Layer where the actual Classification occurs.

## 2. PROBLEM STATEMENT

The project "Object Detection System using Machine Learning Technique" detects objects efficiently based on YOLO algorithm and apply the algorithm on image data and video data to detect objects.

## 3. ARCHITECTURE OF THE PROPOSED MODEL

The Fig-1 shows the Architecture Diagram of the Proposed YOLO Model. Images are given as the input to the system. If Video can also be taken as input as it is nothing but a stream of images. As the name suggests You Only Look Once, the input goes through the network only once and the result of detected object with Bounding Boxes and Labels are obtained.

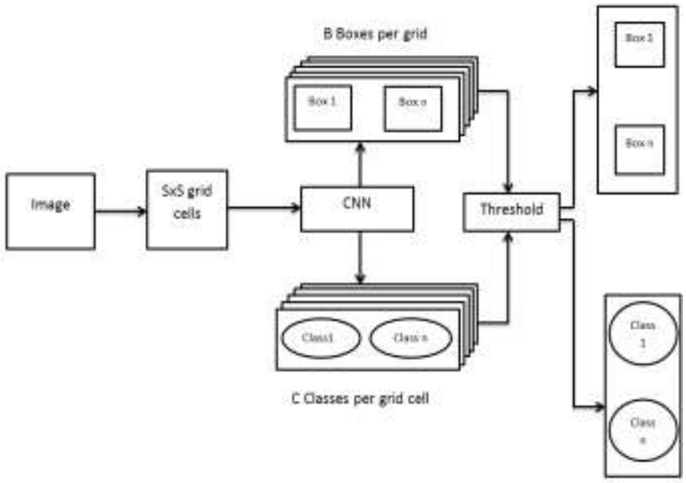


**Fig -1**: YOLO Architecture

The images are divided into SXS grid cells before sending to the Convolutional Neural Network (CNN). B Bounding boxes per grid are generated around all the detected objects in the image as the result of the Convolutional Neural Network. On the other hand, the Classes to which the objects belong is also classified by the Convolutional Neural Network, giving C Classes per grid. Then a threshold is set to the Object Detection. In this project we have given a Threshold of 0.3.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 05 | May 2019 www.irjet.net p-ISSN: 2395-0072

Lesser the Threshold value, more number of bounding boxes will appear in the output resulting in the clumsy output.

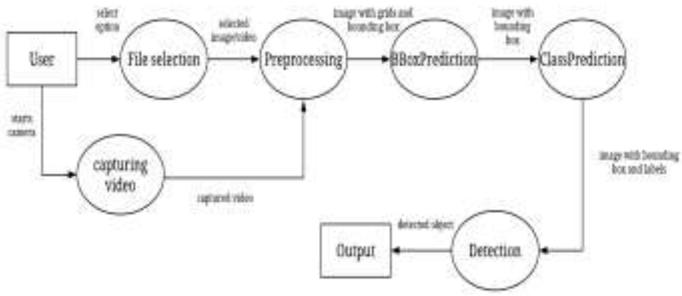


**Fig -2**: Data Flow Diagram of the System

The Fig-2 illustrates the Flow of data in the System. Initially User will be given the options to choose the type of the File to be given to the System as an input. Thus, User can either choose option of File Selection or start the Camera. In the former, User can choose either Image File or a Video File and, in the latter, User can start the Camera module. Once the input is selected Preprocessing is done, where the SXS grids are formed. The resultant thus formed with the grids is send to the Bounding Box Prediction process where the Bounding Boxes are drawn around the detected objects. Next the result from the previous process is sent to the Class Prediction where the Class of the object to which it belongs is predicted. Then it is sent to the detection process where a Threshold is set in order to reduce clumsiness in the output with many Bounding Boxes and Labels in the final Output. At the end an image or a stream of images are generated for image and video or camera input respectively with Bounding Boxes and Labels are obtained as the Output.

## 4. IMPLEMENTATION

This chapter describes the methodology for implementing this project. Following is the algorithm for detecting the object in the Object Detection System.

### 4.1 Algorithm for Object Detection System

1. The input image is divided into SxS grid

2. For each cell it predicts B bounding boxes Each bounding box contains five elements:
   (x, y, w, h) and a box confidence score

3. YOLO detects one object per grid cell only regardless of the number bounding boxes

4. It predicts C conditional class probabilities
5. If no objects exists then confidence score is zero Else confidence score should be greater or equal to threshold value
6. YOLO then draws bounding box around the detected objects and predicts the class to which the object belongs

## 5. RESULTS AND ANALYSIS

This chapter describes the results obtained by the System, different Test Cases used while Testing the System. We used pretrained dataset of COCO which had 80 classes. The reason why 80 classes because a greater number of classes resulted in the overfitting of the data. Following section will describe the different Test Cases and the results obtained.

### 5.1 Test Cases

The Table-1 shows the different Test Cases, the Expected as well as the Test Result.

**Table -1:** Test Cases with Results

| Test Case ID | Test Conditions | Expected Result | Test Results |
|---|---|---|---|
| TC1 | When image is chosen as input | Image with bounding box around the objects and predicted class | SUCCESSFUL |
| TC2 | When video is chosen as input | Video with bounding box around the objects and predicted class | SUCCESSFUL |
| TC3 | When camera is chosen as input | Objects detected in the real time with bounding box, confidence score and predicted class | SUCCESSFUL |
| TC4 | When black and white image is taken as input | Image with bounding box around the objects and predicted class | SUCCESSFUL |
| TC5 | Image with far objects is taken as input | Image with detected objects | UNSUCCESSFUL |
| TC6 | When image with overlapping objects is | Image with bounding box around the objects and | SUCCESSFUL |

| | | | |
|---|---|---|---|
| | taken as input | predicted class | |
| TC7 | When image with far objects is taken as input | Image with detected objects | UNSUCCESSFUL |

### 5.2 Results

This section describes different results obtained by giving various Test Cases described above.

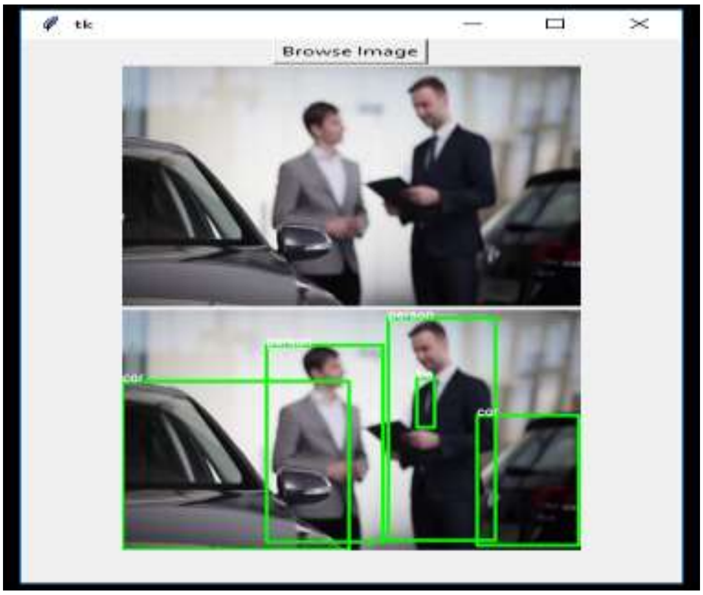


**Fig -3**: Image with Detected Object

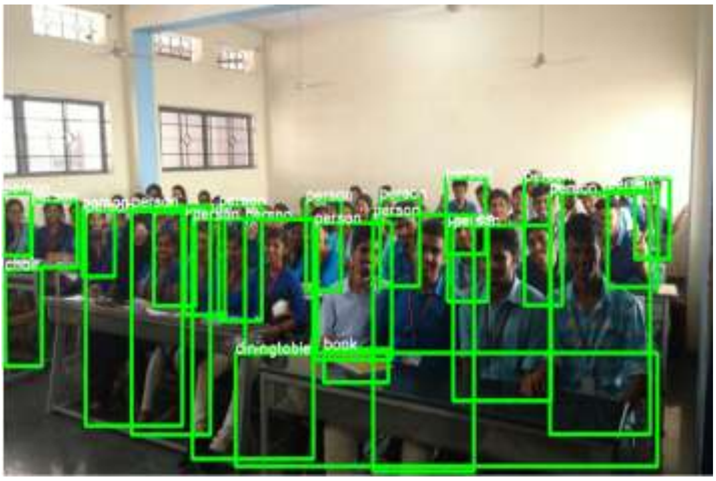

**Fig -4**: Image with Overlapping Objects

The Fig-3 illustrates the output of the Object Detection System. Bounding Boxes are drawn around the Objects detected. Fig-4 illustrates the output obtained when objects are overlapping. This shows that partially visible objects will also be detected by drawing bounding box around it along with the label indicating the class to which it belongs. In the Fig-4 some people are partially visible in the image of a crowded classroom. The system is able to detect every person visible in the image.


**Fig -5**: Output obtained with Video Input


**Fig -6**: Output obtained in Real-Time

The output generated when Video is given as the input is shown in Fig-5. The video that is to be given as input to the system should be in .avi format. Fig-6 illustrates the output when Camera is used to detect the object. Fig-7 illustrates the output generated when a blur image is given as the input. Random bounding boxes are drawn with no detected object. This is one of the drawbacks of the project which given unsuccessful test result.

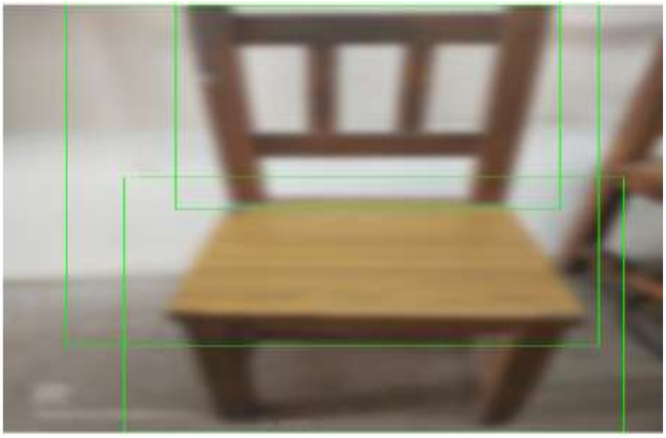

**Fig -7**: Output obtained in Real-Time

## 3. CONCLUSIONS AND FUTURE WORK

The project is developed with objective of detecting real time objects in image, video and camera. Bounding Boxes are drawn around the detected objects along with the label indicating the class to which the object belongs. We have used CPU for the processing in the project. Future enhancements can be focused by implementing the project on the system having GPU for faster results and better accuracy

## REFERENCES

[1] Lele Xie, Tasweer Ahmad, Lianwen Jin , Yuliang Liu, and Sheng Zhang ,“ A New CNN-Based Method for Multi-Directional Car License Plate Detection”, IEEE Transactions on Intelligent Transportation Systems, ISSN (e): 1524-9050, Vol-19, Issue-02, Year-2018, pp. 507-517.

[2] L. Carminati, J. Benois-Pineau and C. Jennewein, “Knowledge-Based Supervised Learning Methods in a Classical Problem of Video Object Tracking”, 2006 International Conference on Image Processing, Atlanta, GA, USA, ISSN (e): 2381-8549, year-2006.

[3] Jinsu Lee, Junseong Bang and Seong-II Yang, “Object Detection with Sliding Window in Images including Multiple Similar Object”, 2017 IEEE International Conference on Information and Communication Technology Convergence (ICTC), Jeju, South Korea, ISBN (e): 978-1-5090-4032-2, December-2017.

[4] Qichang Hu, Sakrapee Paisitkriangkrai, Chunhua Shen, Anton van den Hengel and Faith Porikli,“Fast Detection of Multiple Objects in Traffic Scenes with Common Detection Framework”, IEEE Transactions on Intelligent Transportation Systems, ISSN (e): 1558-0016, Vol-17, Issue-04, Year-2016, pp. 1002-1014.

[5] Haihui Xie, Quingxiang Wu and Binshu Chen, “Vehicle Detection in Open Parks Using a Convolutional Neural Network”, 20015 6th International Conference on Intelligent Systems Design and Engineering Applications(ISDEA), Guiyang, China, ISSN (e): 978-1-4673-9393-5, August-2015.

[6] Malay Shah and Rupal Kapdi, “Object Detection using Deep Learning Networks”, 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, ISBN (e): 978-1-5386-2745-7, June-2017.

[7] Shoji Kido, Yasusi Hirano and Noriaki Hashimoto, “Detection and classification of lung abnormalities by use of convolutional neural network(CNN) and regions with CNN features(R-CNN)”, 2018 International Workshop on Advanced Image Technology, Chiang Mai, Thailand, ISBN (e): 978-1-5386-2615-3, December-2017.

[8] Kai Han, Muyi Sun, Xiaoguang Zhou, Gaunhong Zhang, Hao Dang and Zhicai Liu, “A New Method in Wheel Hub Surface Defect Detection Algorithm Based on Deep Learning”, 2017 IEEE International Conference on Advanced Mechatronic Systems, Xiamen, China, ISBN (e): 978-1-5386-2602-3, December-2017.

[9] Syed Mazhar Abbas and Dr. Shailendra Narayan Singh, “Region based Object Detection and Classification using Faster R-CNN”, 2018 4th International Conference on Computational Intelligence and Communication Technology(CICT), Ghaziabad, India, ISBN (e): 978-1-5386-0886-9, February-2017.

[10] Jinzhuo Wang, Wenmin Wang and Wen Gao, “Multiscale Deep Alternative Neural Network for Large Scale Video Classification”, IEEE Transactions on Multimedia, ISSN (e): 1941-0077, Vol-20, Issue-10, Oct-2018, pp. 2578-2592.

[11] Widodo Budiharto, Alexander A S Gunawan, Jarot S. Suroso, Andry Chowanda, Aurello Patrik and Gaudi Utama, “Fast Object Detection for Quadcopter Drone using Deep Learning”, 2018 3rd International Conference on Computer and Communication System(ICCCS), Nagoya, Japan, ISBN (e): 978-1-5386-6349-3, April-2017.

[12] Pushkar Shukla, Beena Rautela and Ankush Mittal, “A Computer Vision Framework for Automatic Description of Indian Monuments”, 2017 13th International Conference on Signal Image Technology and Internet-Based Systems(SITIS), Jaipur, India, ISBN (e): 978-1-5386-4283-2, December-201.

[13] M. Buric, M. Pobar and Ivasic-Kos, “Object Detection in Sports Videos”, 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics(MIPRO), Opatija, Croatia, ISBN (e): 978-953-233-095-3, May-2018.

[14] Yin-Lon Lin, Yu-Min Chiang and Hsiang-Chen Hsu, “Capacitor Detection in PCB Using YOLO Algorithm”, 2018 International Conference on System Science and Engineering(ICSSE), New Taipei City, Taiwan, ISBN (e): 978-1-5386-6285-4, December-2017

[15] Romain Vial, Hongyuan Zhu, Yonghong Tian and Shijian Lu “Search Video Action Proposal with Recurrent and Static YOLO”,2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, ISBN (e): 978-1-5090-2175-8, September-2017.