INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 05 ISSUE: 07 | JULY - 2021 ISSN: 2582-3930

Real Time Object Detection and Recognition Using Artificial Neural Network and Deep Learning

Mr. Santhosh,Jayesh Shetty,Fathimath Hafeefa,Pranav S Khodanpur

Dep. of Computer Science, Srinivas University College of Engineering & Technology
Mukka, Mangalore, 574146

Abstract - The world of computing has changed a lot in recent years. It's not a long time ago when image recognition and artificial intelligence were considered as idea from a science-fiction novel.the internet age has given rise to a new generation of data scientist who analyze the vast data collect and generate prediction algorithm with the help of data.it is estimated that approximately 300 million images are uploaded to the internet every single day.This enormous collection of data lead to the birth of a new field of computer science called deep learning.Where we use sample data to generate meaning results. Neural networking is another field that has seen a surge in popularity because of the vast amount of data collected thanks to the rise in popularity of the internet. Artificial Neural Networks(ANN) is being used for image processing in various fields of activity such as geotechnics,civil engineering,mechanics,industrial surveillance, defence and transport. Main objective is detection of objects using You Only Look Once (YOLO) algorithm, this method having advantages while compared to other algorithms. YOLO completely considers images by predicting boxes , finding probabilities for boxes and also detects the image faster than other algorithms.

Key Words: Image Detection,YOLO,COCO Dataset and Artificial Neural Network.

I. INTRODUCTION

Recognizing objects from images and videos is called Object Detection. The computer's vision through cameras connected (real time) or through the existing pictorial data is processed to differentiate various elements of the picture or frame. Object detection is considered to be most important as it helps to perform various tasks by automation without the involvement of humans, hence we should use highly accurate algorithms. Various Machine or Deep Learning algorithms to detect or recognize images or frames are Regional Convolution Neural(R-CNN's), You Only Look Once(YOLO), Support Vector Machine(SVM),Convolution Neural Network(CNN). This thesis describes the procedure or process in which objects are detected or recognized from real time video or real time picture frames. Like an image or a video from the marketplace where there are people,shops,animals,vehicles etc this distinguished data can be further used to wide purpose using YOLO Algorithm.

II. Artificial Neural Network

Artificial Neural Network is a way to replicate how a human brain works and analyses things and situations. A human brain has neurons that send signals to the brain about what it has to do. A neuron can only be in one state at a time it can either send signals or not it can be represented in computer language as sending a 0 or 1. Many years of research has led to the development of the Artificial Neural Network which is an algorithm that replicates a human brain and makes it possible for the computer to make decisions faster. It contains multiple Neurons that send individual signals to a program code or a function to perform a particular set of tasks.

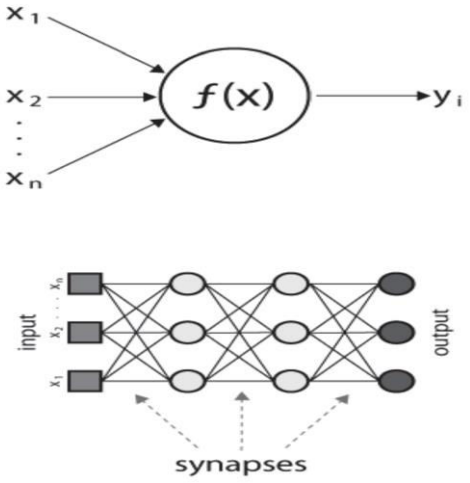

figure 1:Representation of Artificial Neural Network

III. LITERATURE SURVEY

Real time object detection is a very active field where many papers have addressed neural computations, however much of work is yet to be conducted to increase efficiency, and real time object detection needs to be brought into day to day lives of people to make their living better .

[1]Objects are detected and tracked in multiple regions using a single fixed camera view in real time for security related applications.

[2] Deval Jansari The inspiration of the proposed strategy is to take

care of regular issues for various object tracking like changes in the background, blur due to the motion, occlusion, movement of object, and overlapping.This study suggests colour-based likelihood coordination for ongoing objects, recognize motion of objects and trace the same objects which show up in the consequent casing. At first, recognition is conveyed by utilizing two extraordinary methods (i.e. Background Subtraction Modelling and Optical Flow Method). Advantages and disadvantages of the two strategies are described. At that point,objects in the resulting frame are exclusively distinguished by colour-based probability.

[3] Jorge García in the study of bidirectional object detection that is based on the content of the stereo system,and importance of camera placement,distance estimation and K-means Algorithms were discussed.

IV. EXISTING SYSTEM

The current system of image detection only works with static images and cannot be used to detect objects from a video. The current system is very graphic intensive and requires a high amount of processing power and the speed is not at power either. The accuracy of the current system is not that great either. Withe development of the fields like big data and increase in the amount of data that is generated everyday we can use that data to improve the system and make it better.

V. PROPOSED SYSTEM

A. YOLO

YOLO is a convolutional neural network (CNN) for object detection in real-time. YOLO stands for You Only Look Once as the name suggests this algorithm only looks at the image once and detects all the objects. It applies a convolutional neural network to the entire image at once and makes predictions. It divides the image into blocks and based on those blocks it determines which block section contains an object or the section of the image that it thinks is relevant. It creates a bounding box around that object and tries to detect what the object is and makes predictions based on the dataset we have used to train the model. We are using images for a general reposty to detect normal day today objects.

B. COCO Dataset

The Coco dataset is a dataset of images developed by many non profit organizations. It is a very helpful repository of images that can be used to train any image recognition algorithm.

C. Training Algorithm

The Algorithm is trained and tested using the COCO dataset to check the accuracy and speed of processing and all the necessary changes are done to improve the accuracy of the system. A custom dataset is then generated to check if there is any change in performance. Once all the training and testing is done the Algorithm can be deployed

VI. HARDWARE REQUIREMENTS

•Processor: Intel-core, i-3 / i-5 / i-7
•Speed: 2.0GHz
•Hard Disk: 1 TB
•I/O: Basic keyboard and mouse.
•RAM: 4 GB
•Monitor: SVGA

VII. SOFTWARE REQUIREMENTS

•Programming Language: Python
•Operating system: Windows 7 or above.
•Software: Pycharm

VIII. ARCHITECTURAL DESIGN

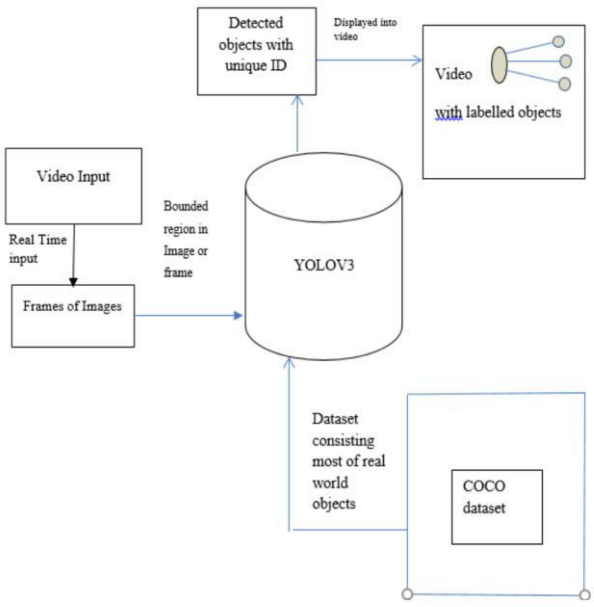

Figure 2:Architecture Design

IX. SYSTEM IMPLEMENTATION

i. Real time video is collected from the webcam of a laptop or any image capturing device. It depends on what the user wants to process; the system works with both a video input or an image input.
ii. Once the user decides what kind of input he or she wants to provide the input is then processed.
iii. If the input provided is an image the processing takes place instentently. If it is a video input then the video is converted into frames with bounded regions.
iv. The bounded frames are then processed by the YOLO V3 algorithm to detect the objects present in the frame.
v. The objects are classified based on the dataset provided to the algorithm in our cases it will be processed using the coco dataset and some custom dataset that we created.
vi. Detected Objects are given a unique id and if the same objects are present multiple times they are detected as different objects.
vii. The detected objects have a bounding box around then so that the user can distinguish between the objects and it is easier to recognize.

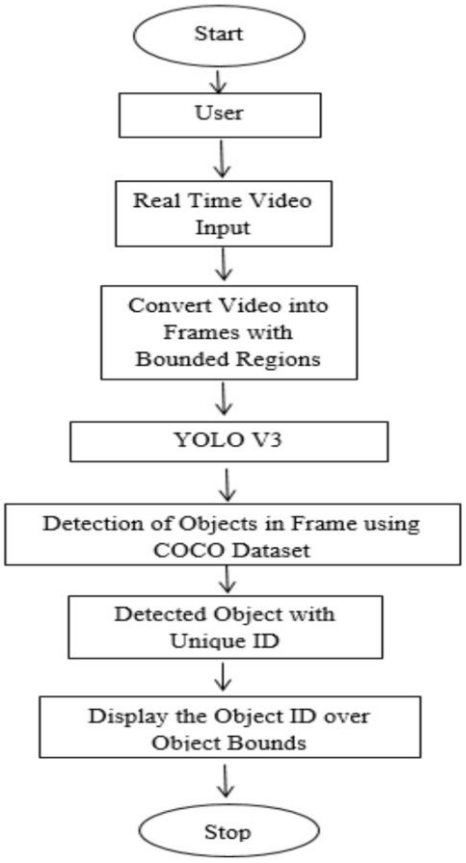

Figure 3:Workflow of the project implementation

CONCLUSION

Real Time Object Detection using the Artificial Neural Networks field has been advanced to a level where now the results are 100% efficient in many cases and an average of above 90% is recorded for detecting real world objects present around us.
In this project most accurate technique has been identified for Real time object detection that increases the efficacy rates in above results above 90% that was obtained from live stream of video input from web camera using YOLOv3 algorithm, pre-processing for faster detection was implemented successfully in the project to support real time object detection. Results obtained from the project can be further used in the field of applied real time object detection that could help blind people as automated real time guide and other such similar applications.

REFERENCES

[1] M. B. Blaschko and C. H. Lampert. Learning to localize objects with structured output regression. In Computer Vision– ECCV 2008, pages 2–15. Springer, 2008.

[2] D. Jansari, S. Parmar and G. Saha, "Real-time object tracking using color-based probability matching," IEEE International Conference on Signal Processing, Computing and Control (ISPCC), Solan, 2013, pp. 1-6.

[3] J. García, A. Gardel, I. Bravo, J. L. Lázaro and M. Martínez, "Tracking People Motion Based on Extended Condensation Algorithm," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 43, no. 3, pp. 606-618, May 2013.

[4] E. Maggio and A. Cavallaro, "Learning Scene Context for Multiple Object Tracking," IEEE Transactions on Image Processing, vol. 18, no. 8, pp. 1873-1884, Aug. 2009.

[5] Mukesh Tiwari, Rakesh Singhai, "A Review of Detection and Tracking of Object from Image and Video Sequences", International Journal of Computational Intelligence Research, Volume 13, Number 5 (2017), pp. 745-765.

[6] J. Zhu, Y. Lao and Y. F. Zheng, "Object Tracking in Structured Environments for Video Surveillance Applications," IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 2, pp. 223-235, Feb. 2010.

[7]Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi," You Only Look Once: Unified, Real-Time Object Detection " 2016 IEEE Conference on Computer Vision and Pattern Recognition.

[8] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 2155–2162. IEEE, 2014.

[9] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. International Journal of Computer Vision, 111(1):98–136, Jan. 2015.