IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11790

Object Tracking and Counting using Computer Vision and Deep Learning

Prashant Abbi[1], Senthooran B[2], Prof Merin Meleet[3]

Student, Information Science and Engineering, RV College of Engineering, Bengaluru, India[1]

Student, Information Science and Engineering, RV College of Engineering, Bengaluru, India[2]

Assistant Professor, Information Science and Engineering, RV College of Engineering, Bengaluru, India[3]

Abstract: Computer Vision is an important field that has been revolutionised by the emergence of Deep Learning and Neural Networks. Object Detection is an important subclass of Object Detection that involves image classification and object localization. There are 2 major classes of object detection algorithms - one stage detectors and multi-stage detectors. Multi-stage detectors like Region-based Convolutional Neural Networks (R-CNN), Fast R-CNN and Faster-RCNN first make region proposals and then make separate predictions for each of these regions. Single stage detectors like YOLO (You Only Look Once) require only one single pass through the convolutional network and predict the bounding boxes in one go. Single shot detectors like YOLO perform better when speed is the most important factor, even more so than accuracy. YOLO has applications in real-time systems like autonomous driving, crowd management, etc. The algorithm developed performs object counting in addition to object detection.

Keywords: YOLO, Object Detection, Object Counting, Convolutional Neural Networks.

I. INTRODUCTION

Computer Vision is seeing great interest in the past few years with the advent of Artificial Intelligence and Machine Learning as a part of the Fourth Industrial Revolution. Object Detection can be considered a sub-domain of Computer Vision. It has applications in various fields, some of which are surveillance, defence, healthcare, and traffic analysis.
On a certain level, Computer vision is all about pattern recognition. So, one way to train a computer how to understand visual data is to feed it lots of images that have been labelled, and then subject them to algorithms that allow the computer to hunt down patterns in all the elements that relate to those labels.
Before the advent of deep learning, the tasks that computer vision could perform were extremely limited and required a lot of manual coding and effort by developers and human operators.

Deep learning provided a fundamentally different approach to doing machine learning. Deep learning relies on neural networks, a general-purpose function that can solve any problem representable through examples. When you provide a neural network with many labelled examples of a specific kind of data, it will be able to extract common patterns between those examples and transform it into a mathematical equation that will help classify future pieces of information. Deep Learning plays a significant role in object detection and tracking as it has the capability to improve with increasing subject data. Most current computer vision applications such as cancer detection, self-driving cars and facial recognition make use of deep learning.

Object recognition is a general term to describe a collection of related computer vision tasks that involve identifying objects in digital photographs. Image classification involves predicting the class of one object in an image. Object localization refers to identifying the location of one or more objects in an image and drawing a bounding box around their extent. Object detection combines these two tasks and localizes and classifies one or more objects in an image.
Object Tracking and Counting involves detecting objects in an image or video and tracking their movement within the image or frame limits and counting the number of instances of each object class being tracked. Object Tracking and Counting can have important applications in traffic-flow analysis and surveillance systems.

II. METHODOLOGY

The proposed approach utilises the YOLO algorithm for object detection. The YOLO Algorithm was trained over the COCO (Common Objects in Context) dataset. It uses neural networks to provide real-time object detection.
The developed system consists of 3 modules - the Video Module, the Detection Module, and the Counting Module.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11790

The video module is responsible for processing the video or feed before object detection can be performed. It allows us to choose the video or the camera stream on which detection and counting is to be performedIn case of a video- the path to the video is specified, and in case of a live camera stream, the url to its feed is fed as input. OpenCV is the main library employed for the task of processing the video. OpenCV (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. OpenCV helps us to break the video into individual frames which can be passed to the detection module for detecting objects in real time.

The Object Detection Module uses the YOLO algorithm. YOLO employs a CNN to perform detections in a single pass through the network. This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously. YOLO makes use of 3 processes for object detection - Residual blocks, Bounding Box Regression and Intersection Over Union.

- In Residual blocks, the image is divided into cells in a grid pattern. The grid cells can then detect objects that appear within them. YOLO uses a single bounding box regression to predict the height, width, centre, and class of objects. This means that detections are solved as a regression problem rather than a classification problem.
- A bounding box is a rectangle that encloses the detected object. Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap.
- YOLO uses IOU to provide an output box that surrounds the objects perfectly. Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

The final step is counting. The Detection Module gives us frames with object detection carried out and these are passed to the final module - the Counting Module. The Counting Module counts the various bounding boxes of the detected object classes and outputs the number to the screen.

III. SIMULATION AND EXECUTION

Based on the methodology discussed in previous sections, a python program was developed making use of OpenCV, TensorFlow, numpy, Pandas and PIL (Python Image Library)
Various object detections were trained in this model. The following results were obtained after detection, tracking and counting on video feeds and camera streams.

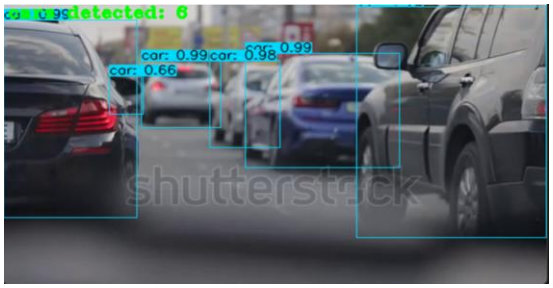

Fig. 1.

Fig. 2.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11790

IV. CONCLUSION

A system for detecting the various COCO classes was developed. The system also has the capability of tracking the detected objects as they stay in the field of vision. An additional feature of counting the number of instances of the various classes is also included in the project. This project has various applications in the fields of traffic analysis, law enforcement, crowd management, etc.

Traffic police can use the application to manage traffic more effectively. This has the potential to better manage traffic and make cities less congested and allow for lower commuting times.

Law Enforcement can use person detection to track anti-social elements more easily to nail down their location easily.

Crowd Management applications can be utilized to manage crowds in public spaces such as railway stations, airports, and bus stations.

REFERENCES

[1]. H. H. Nguyen, T. N. Ta, N. C. Nguyen, V. T. Bui, H. M. Pham and D. M. Nguyen, "YOLO Based Real-Time Human Detection for Smart Video Surveillance at the Edge," 2020 IEEE Eighth International Conference on Communications and Electronics (ICCE), 2021, pp. 439-444, doi: 10.1109/ICCE48956.2021.9352144.

[2]. Preetham T.P., Sanjitha Gowda, Dr. Mamatha G.S, Prof. Anisha B.S International Journal for Research in Applied Science and Engineering Technology IJRASET, 2020. ISSN – 2321-9653.

[3]. L. R. Cerna, G. Cámara-Chávez, D. Menotti, Computer Science Department, Federal University of Ouro PretoOuro Preto, MG, Brazil.

[4]. Nirmal Kumar Saraswat, Arvinda Kushwaha. International Journal of Engineering and Advanced Technology (IJEAT)ISSN: 2249- 8958, Volume-8 Issue-6, August 2019

[5]. Michael Bryan, P. Lumaban, Giged T. Battung. International Journal of Engineering Research & Technology (IJERT). ISSN: 2278-0181. Vol. 10 Issue 08, August-2021

[6]. Shaokang Chen, Erik Berglund, Abbas Bigdeli, Conrad Sanderson, Brian C. Lovell. 2008 IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance, 2008

[7]. Ullah, Rehmat & Hayat, Hassan & Siddiqui, Afsah & Siddiqui, Waleed & Islam, Muhammad & Karami, Ghulam. (2022). Mathematical Problems in Engineering. 2022. 10.1155/2022/3276704.

[8]. G. N. Made Kris Raya, A. N. Jati and R. E. Saputra, 2017. International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics), 2017, pp. 1-5, doi: 10.1109/ROBIONETICS.2017.8203427.

[9]. SREEKANTHA.B, VIJAYALAKSHMI S A, HARI PRASAD B, HRITHIK C, NAGAJYOTHI M S, RAKSHITHA S. International Research Journal of Engineering and Technology. e-ISSN: 2395-0056. Volume: 08 Issue: 06 | June 2021

[10]. Fung, C.C. and Jerrat, N. (2000). Published in: 2000 TENCON Proceedings. Intelligent Systems and Technologies for the New Millennium (Cat. No.00CH37119)

[11]. Shengyu Lu, Beizhan Wang, Hongji Wang, Lihao Chen, Ma Linjian, Xiaoyan Zhang, Computers & Electrical Engineering, Volume 77, 2019, Pages 398-408, ISSN 0045-7906

[12]. W. Yang and Z. Jiachun, "Real-time face detection based on YOLO," 2018 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII), 2018, pp. 221-224, doi: 10.1109/ICKII.2018.8569109.