Volume 10, Issue 1, January 2023 ISSN: 2348-8565 (Online)
International Journal of Modern Engineering and Research Technology
Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

Object Detection and Tracking Using Machine Learning Algorithm

K. Venkatesh
Research Scholar
Computer Science and Engineering
Annamalai University
Annamalainagar (T.N), India
Email: venkiur@gmail.com

S. Pasupathy
Associate Professor
Department of Computer Science and Engineering
Annamalai University
Annamalainagar (T.N), India

ABSTRACT

Deep learning in machine learning has gained a tremendous influence on how the world is adapting to Artificial Intelligence since past few years. Some of the popular object detection algorithms are Region based Convolutional Neural Networks (RCNN), Faster-RCNN, Single Shot Detector (SSD) and You Only Look Once (YOLO). Amongst these, Faster-RCNN and SSD have better accuracy, while YOLO performs better when speed is given preference over accuracy. Deep learning combines SSD and Mobile Nets to perform efficient implementation of detection and tracking. This algorithm performs efficient object detection while not compromising on the performance. Although there are number of object detection systems implemented in past researchers, there still remains a constant demand for new, better and accurate detection systems. Current detection systems make use of classifiers to perform detection. We are implementing a machine learning model which can detect objects using the concept of convolutional neural network (CNNs) using yolo algorithm.

Keywords:— *YOLO, SSD (Single Shot Detector), RCNN (Region based Convolutional Neural Networks), Convolutional Neural Network (CNNs)*

I. INTRODUCTION

Object detection and location in digital images has become one of the most important applications for industries to ease user, save time and to achieve parallelism. This is not a new technique but improvement in object detection is still required in order to achieve the targeted objective more efficiently and accurately. Object detection involves detecting region of interest of object from given class of image. Different methods are –Frame differencing, Optical flow, Background subtraction. This is a method of detecting and locating an object which is in motion with the help of a camera.

The main purpose of object detection and tracking is to identify and locate one or more effective targets from still image or video data. Object recognition and tracking reduces human efforts and provides efficiency. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval, traffic analysis and video surveillance. Object tracking has many practical applications including surveillance, medical imaging, traffic flow analysis, self-driving cars, people counting and audience flow analysis, and human-computer interaction. Technically, object tracking starts with

object detection identifying objects in an image and assigning them bounding boxes. YOLO struggles to detect and segregate small objects in images that appear in groups, as each grid is constrained to detect only a single object. Small objects that naturally come in groups, such as a line of ants, are therefore hard for YOLO to detect and localize.

II. Literature Review

Object Detection and Recognition in Images

Object Recognition is a technology in the field of computer vision. It is considered to be one of the difficult and challenging tasks in computer vision. Many approaches have been proposed in the past, and a model with a new approach which is not only fast but also reliable. Easynet model has been compared with various other models as well. Easynet model looks at the whole image at test time so its predictions are informed by global context. At the prediction time; our model generates scores for the presence of the object in a particular category. It makes predictions with a Single network evaluation. Here object detection is a regression problem to spatially separated bounding boxes and associated class probabilities.[1]

Deep Neural Networks for Object Detection

Deep Neural Networks (DNNs) have recently shown outstanding performance on image classification tasks. In this paper we go one step further and address the problem of object detection using DNNs, that is not only classifying but also precisely localizing objects of various classes. We present a simple and yet powerful formulation of object detection as a regression problem to object bounding box masks. We define a multi-scale inference procedure which is able to produce high-resolution object detections at a low cost by a few network applications. State-of-the-art performance of the approach is shown on Pascal VOC. [2]

Computer Vision-Based Object Detection and Classification

Computer vision techniques become particularly important in agriculture applications due to their fast response, high accuracy and strong adaptability. Two of the most demanding and widely studied applications relate to object detection and classification. The task is challenging due to variations in product quality differences under certain complicate circumstances influenced by nature and human. Research in these fields has resulted in a wealth of processing and analysis methods. In this paper, we explicitly explore current advances in the field of object detecting and categorizing based on computer vision, and a comparison of these methods is given.[3]

Daily Objects Detection Based on Deep Neural Network

With the rapid development of deep learning, great breakthroughs have been made in the field of object detection. In this article, the deep learning algorithm is applied to the detection of daily objects, and some progress has been made in this direction. Compared with traditional object detection methods, the daily objects detection method based on deep learning is faster and more accurate. The main research work of this article: 1. collect a small data set of daily objects; 2. in the TensorFlow framework to build different models of object detection, and use this data set training model; 3. the training process and effect of the model are improved by fine-tuning the model parameters.[4]

Viola Jones algorithm:

Despite being an outdated framework, Viola -Jones is quite powerful, and its application has proven to be exceptionally notable in real-time face detection. This algorithm is painfully slow to train but can detect faces in real-time with impressive speed. Given an image (this algorithm works on grayscale image), the algorithm looks at many smaller subregions and tries to find a face by looking for specific features in each subregion. It needs to check many different positions and scales because an image can contain many faces of various sizes. Viola and Jones used Haar-like features to detect faces in this algorithm.[5]

III. Work Description

Existing System

Many problems in computer vision were saturating on their accuracy before a decade. However, with the rise of deep learning techniques, the accuracy of these problems drastically improved. One of the major problems was that of image classification, which is defined as predicting the class of the image. There is no object detection in existing system by using OpenCV. The limitations of existing Systems is difficult to classify the moving object tracking and also the accuracy of object detection is very low.

Proposed System

Moving object detection is the initial step for the process of analyzing a video. This is done either in each and every frame or when the object first appears in the video. The classification of objects is done based on their shape features of the motion region. Object tracking is the next step followed by the object detection. It is a technique used to track and also the travelling direction of objects.

Dense Optical flow: These algorithms help estimate the motion vector of every pixel in a video frame.

Sparse optical flow: These algorithms, like the Kanade-Lucas-Tomashi (KLT) feature tracker, track the location of a few feature points in an image.

Kalman Filtering: A very popular signal processing algorithm used to predict the location of a moving object based on prior motion information. One of the early applications of this algorithm was missile guidance! Also as mentioned here, “the on-board computer that guided the descent of the Apollo 11 lunar module to the moon had a Kalman filter”. The advantages of proposed system is to here we can detect the object for uploaded video file. It can easily track the moving object by using point tracking and kernel tracking methods. It not only detects the object but also track the object in the videos.

IV. Methodology

General Architecture

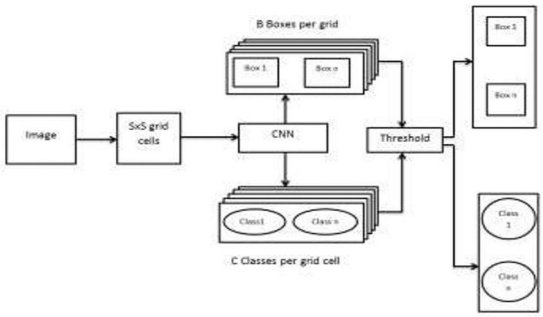

Figure 1: Architecture

The Figure 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.

ALGORITHM

YOLO Algorithm

To understand the YOLO algorithm, it is necessary to establish what is actually being predicted. Ultimately, we aim to predict a class of an object and the bounding box specifying object location. Each bounding box can be described using four descriptors. In addition, we have to predict the pc value, which is the probability that there is an object in the bounding box. center of a bounding box (bxby) width (bw) height (bh), value cis corresponding to a class of an object (e.g., car, traffic lights, etc.). In addition, we have to predict the pc value, which is the probability that there is an object in the bounding box.

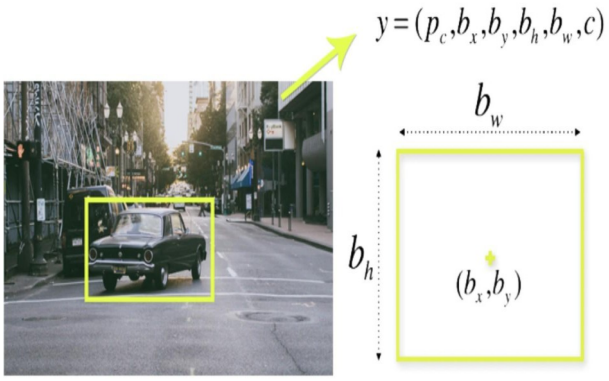

Figure 2: Bounding box probability calculation

As we mentioned above in figure 2, when working with the YOLO algorithm we are not searching for interesting regions in our image that could potentially contain an object. Instead, we are splitting our image into cells, typically using a 19×19 grid. Each cell is responsible for predicting 5 bounding boxes (in case there is more than one object in this cell). Therefore, we arrive at a large number of 1805 bounding boxes for one image. Rather than seizing the day with YOLO and Carpe Diem, we're looking to seize object probability. The exchange of accuracy for more speed isn't reckless behavior, but a necessary requirement for faster real-time object detection.

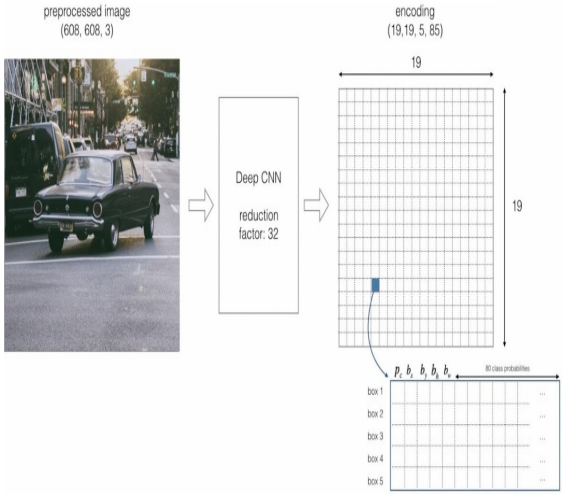

Figure 3: Image size-reduction

From figure 3 Most of these cells and bounding boxes will not contain an object. Therefore, we predict the value pc, which serves to remove boxes with low object probability and bounding boxes with the highest shared area in a process called non-max suppression.

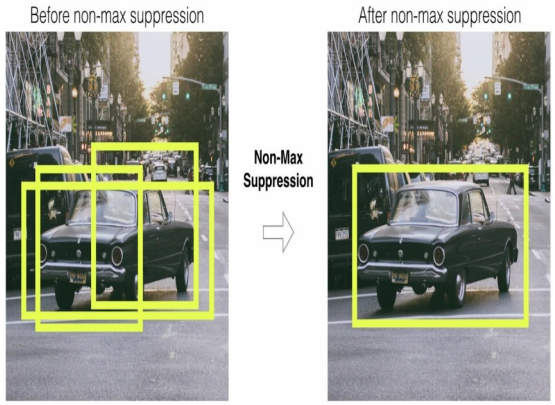

Figure 4: Non-max suppression

From figure 4 it shows the object detected before non-max suppression and after non-max suppression.

V. MODULE DESCRIPTION

The complete paper has been implemented using Python programming language and various modules available in Python. Those modules are acquiring the data. Training using coco data set Detect objects using CNN Using non-max suppression.

Implementation and Testing

Input Design

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considers the following things. It shows what data should be given as input and it shows how the data should be arranged or coded. The dialog to guide the operating personnel in providing input. Methods for preparing input validations and steps to follow when error occur.

Objectives:

Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus, the objective of input design is to create an input layout that is easy to follow

Output Design

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements. Select methods for presenting information. Create document, report, or other formats that contain information produced by the system. The output form of an information system should accomplish one or more of the following objectives. Convey information about past activities, current status or paperions of the Future. Signal important events, opportunities, problems, or warnings. Trigger an action. Confirm an action.

VI. Results and Discussions

Efficiency of the Proposed System

Moving object detection is the initial step for the process of analyzing a video. This is done either in each and every frame or when the object first appears in the video. The classification of objects is done based on their shape features of the motion region. Object tracking is the next step followed by the object detection. It is a technique used to track and also the travelling direction of objects. Here we can detect the object for uploaded video file. It can easily track the moving object by using point tracking and kernel tracking methods. It not only detect the object but also track the object in the videos.

Comparison of Existing and Proposed System

Where as in existing system, Object Detection is a field which is always under re- search and exploration by many programmers and developers around the world. The existing approaches have done a decent job in object recognition. They have collected huge data sets for training their machine learning model. Accuracy was not as satisfactory as everyone would expect. This may be because treating the object detection problem as a classification problem and implementing classification algorithms on the data sets. Where in proposed system, it tries to overcome the drawbacks of the existing approaches. Neural network is an emerging technique in Machine learning which has advanced techniques within itself. Those techniques have been used in this paper. The initial step of the system is to divide the input image into a grid of size 13x13. OpenCV has been used for this step. The further steps involve detection of objects from each cell. Some objects may spread across multiple cells and hence they get detected multiple times. As a result, multiple boxes will be drawn around a single object. To remove the additional boxes and leave one box which best fits the objects, we use an algorithm called Non-Max Suppression Algorithm.80 objects can be detected using this machine learning model. Accuracy has been increased a lot due to the use of YOLOV3 which is a type of neural network.

Output

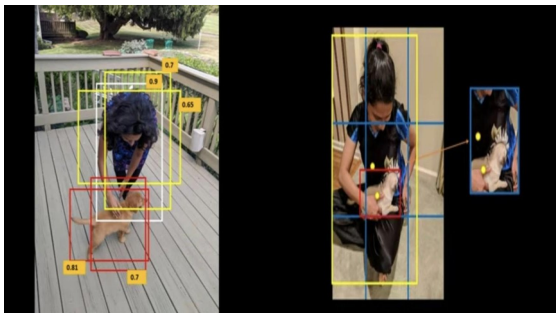

Figure 5: Output 1-Detects two objects from the given input images

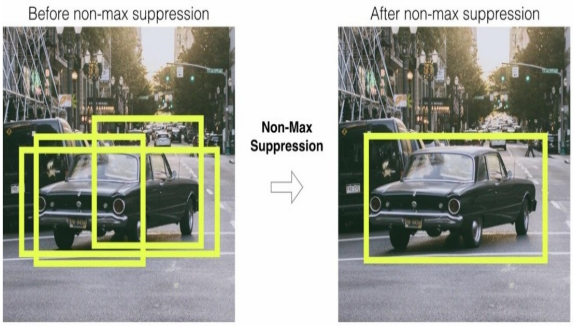

Figure 6: Output 2- Input image before and after non-max-suppression

VII. Conclusion

An accurate and efficient object detection system has been developed which achieves comparable metrics with the existing state-of-the-art system. The concept is well extendable in applications like Intelligent Robots, Automatic Guided Vehicles, Enhancement of Security Systems to detect the suspicious behavior along with detection of weapons, identify the suspicious movements of enemies on boarders with the help of night vision cameras and many such applications. In the

proposed method, background subtraction technique has been used that is simple and fast. For robotic application or automated vehicle assistance system, due to the movement of camera, backgrounds are continuously changing leading to implementation of some different segmentation techniques like single Gaussian mixture or multiple Gaussian mixture models. Object identification task with motion estimation needs to be fast enough to be implemented for the real time system. Still there is a scope for developing faster algorithms for object identification. Such algorithms can be implemented using FPGA or CPLD for fast execution.

REFERENCES:

[1] Sandeep Kumar, Aman Balyan and Manvi Chawla (Computer Science and Engineering Maharaja Surajmal Institute of Technology, New Delhi, India.), 2019.

[2] Christian Szegedy Alexander Toshev Dumitru Erhan, Google, Inc., July 2018.

[3] Juan Wu, Bo Peng, Zhenxiang Huang and Jietao Xie, (College of Information and Electrical Engineering China Agricultural University, Beijing, China. Suzhou Institute for Advanced Study University of Science Technology of China, Suzhou, China), Dec 2019.

[4] Sheng Ding, Kun Zhao. (Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education, Wuhan University of Technology, Wuhan, China; Fiber home Telecommunication Technologies Co.,Ltd, Wuhan, China), 2020.

[5] Wei Liu and Alexander C. Berg, "SSD: Single Shot Multi Box Detector", Google Inc., Dec 2018.

[6] Andrew G. Howard, and Hartwig Adam, "Mobile Nets: Efficient Convolutional Neural Networks for Mobile Vision Applications", Google Inc., 17 Apr 2019.

[7] Justin Lai, Sydney Maples, "Ammunition Detection: Developing a Real Time Gun Detection Classifier", Stanford University, Feb 2020.

[8] Shreyamsh Kamate, "UAV: Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles", Texas A and M University, 2019.

* * * * *