Smart Object Detection Using Machine Learning in Python

Parikshit Jain, Nilesh Wagmode, Gaurav Jalambe and Akanksha Kaldate

May 16, 2023

SMART OBJECT DETECTION USING MACHINE LEARNING IN PYTHON

ABSTRACT:

Efficient and accurate object detection has been an important topic in the advancement of computer vision systems. With the advent of deep learning techniques, the accuracy for object detection has increased drastically. The project aims to incorporate state-of-the-art technique for object detection with the goal of achieving high accuracy with a real-time performance. The network is trained on the most challenging publicly available data-set, on which an object detection challenge is conducted annually. The resulting system is fast and accurate, thus aiding those applications which require object detection.

KEYWORDS:

YOLO, OpenCV, Numpy, Tensorflow, COCO.

INTRODUCTION:

Smart object detection using machine learning is a fascinating and rapidly evolving area of computer vision that has the potential to revolutionize many fields. In simple terms, smart object detection refers to the process of automatically identifying and localizing objects in images or video streams using machine learning algorithms.

This technology has a wide range of practical applications, including surveillance, autonomous vehicles, robotics, and augmented reality. The main goal of smart object detection is to accurately recognize and classify objects in a given image or video, even in challenging environments where the objects may be partially obscured or in motion.

Machine learning algorithms are at the core of smart object detection, and there are various approaches to achieving this task. Some of the most commonly used techniques include supervised learning, unsupervised learning, and deep learning. Deep learning approaches, such as convolutional neural networks (CNNs), have proven to be particularly effective for object detection and classification.

The process of smart object detection involves analysing the features and patterns in the image or video stream, and using this information to make predictions about the location and identity of objects. This is achieved by training the machine learning algorithms on large datasets of labelled images, which enable the algorithms to learn the characteristics of different objects.

One of the main advantages of smart object detection is its ability to perform complex tasks in real-time. For example, smart cameras can be used for real-time surveillance to detect and identify potential threats, such as intruders or suspicious packages. Similarly, smart vehicles can use object detection to navigate complex environments, avoiding obstacles and recognizing traffic signs.

Literature Survey:

Detecting and identifying multiple objects in an image is hard for machines to recognize and classify. However, a noteworthy effort has been carried out in the past years in the detection of objects using convolutional neural networks (CNNs). In the object detection and recognition field, neural networks are in use for a decade but became prominent due to the improvement of hardware new techniques for training these networks on large datasets. In object detection and recognition, researchers have used deep learning for learning features directly from the image pixels, which are more effective than the manual features. Recently deep learning-based algorithms remove the manual features extraction methods and directly use features extracting methods from the original images. ▪is methodology has been successfully proven in feature pyramid network (FPN), single shot detector (SSD), and deconvolutional single shot detector (DSSD). Deep learning is a prevailing direction in the field of machine learning. In, researchers showed that the CNNs inherit the advantages of deep learning, which makes their results in themed of object detection and recognition greatly improved compared with the traditional methods. Researchers had made many efforts to use stochastic gradient descent and backpropagation to train deep networks for object detection. ▪use networks were able to learn but were too slow in practice to be useful in real-time applications; the technique in showed that stochastic gradient descent by back-propagation was effective in training CNNs. CNNs became in use but fell out of fashion due to the support vector machine as in and other simpler methods like linear classifiers as in. New techniques that have been developed recently, show higher image classification accuracy in ImageNet large scale visual recognition. ▪ese techniques have brought much more easiness to train large and deeper networks and shown enhanced performance. Newly, approaches have been established to identify vehicles and other objects from videos or static images using deep convolutional neural networks (DCNN). For example, faster R-CNN proposes candidate regions and uses CNN to confirm candidates as valid objects. YOLO uses end-to-end unified, fully convolutional network structure that predicts the objectless assurance and the bounding boxes concurrently over the whole image. SSD [31] out-performs YOLO by discretizing the production space of bounding boxes into a set of avoidance boxes over different feature ratios and scales per feature map location. YOLO-2 achieves state-of-the-art performance in object detection by improving various aspects of its earlier version. A fully convolutional network is utilized for object detection from three-dimensional (3D) range scan data with LIDAR.A 2D-DBN design is proposed, which uses second-order planes instead of first-order vectors as inputs and uses bi-linear projection for retaining discriminative information to develop the recognition rate. Although DCNN based approaches accomplish the state-of-the-art accuracy of de-taction or classification, these approaches often require intensive calculation and a considerable amount of lifedraining data. ▪rough the past few years, to use deep neural networks economically in real-time applications, a sub-spatial amount of work has been done to report these two problems. In this study, a different modified architecture for object detection is addressed, which is capable of providing high accuracy and speed

METHODOLOGY:

In various fields, there is a necessity to detect the target object and also track them effectively while handling occlusions and other included complexities. Many researchers (Almeida and Gutting 2004, Hsiao-Ping Tsai 2011, Nicolas Papadakis and Auer lie Bureau 2010) attempted for various approaches in object tracking. The nature of the techniques largely depends on the application domain. Some of the research works which made the evolution to proposed work in the field of object tracking is depicted as follows.

OBJECT DETECTION:

Object detection is an important task, yet challenging vision task. It is a critical part of many
applications such as image search, image auto-annotation and scene understanding, object tracking.
Moving object tracking of video image sequences was one of the most important subjects in
computer vision. It had already been applied in many computers vision fields, such as smart video
surveillance (Arun Hamper 2005), artificial intelligence, military guidance, safety detection and
robot navigation, medical and biological application. In recent years, a number of successful
single-object tracking system appeared, but in the presence of several objects, object detection
becomes difficult and when objects are fully or partially occluded, they are obtruded from the human
vision which further increases the problem of detection. Decreasing illumination and acquisition
angle. The proposed MLP based object tracking system is made robust by an optimum selection of
unique features and also by implementing the Ad boost strong classification method.

Existing system:

There has been a lot of work in object detection using traditional computer vision techniques (sliding windows, deform-able part models). However, they lack the accuracy of deep learning-based techniques. Among the deep learning-based techniques, two broad class of methods are prevalent: two stage detection

(RCNN, Fast RCNN, Faster RCNN) and unified detection (Yola, SSD).

GAPS IN EXISTING SYSTEM:

Disadvantages of Existing system

1. Less prediction rates
2. less accuracy

Proposed System:

Here we are proposed YOLOV3 and YOLOV3-TINY models, one of the important fields of Artificial Intelligence is Computer Vision the science of computers and software systems that can recognize and understand images and scenes. Computer Vision is also composed of various aspects such as image recognition, object detection, image generation, image super-resolution and more. Object detection is probably the most profound aspect of computer vision due the number of practical use cases.

Object detection refers to the capability of software systems to locate objects in an image/scene and identify each object. It has been widely used for face detection, vehicle detection, pedestrian counting, web images, security systems and driver less cars. There are many ways object detection can be used as well in many fields of practice. Like every other computer technology, a wide range of creative and amazing uses of object detection will definitely come from the efforts of computer programmers and software developers. Getting to use modern object detection methods in applications and systems, as well as building new applications based on these methods is not a straight forward task.

Early implementations of object detection involved the use of classical algorithms, the popular computer vision library. However, these classical algorithms could not achieve enough performance to work under different conditions.

Advantages of proposed system

1. High Accuracy
2. Very Effective Models

CONCLUSION

An accurate and efficient object detection system has been developed which achieves comparable metrics with the existing state-of-the-art system. This project uses recent techniques in the field of computer vision and deep learning. Custom data set was created using labelling and the evaluation was consistent. This can be used in real-time applications which require object detection for Pre-processing in their pipeline. An important scope would be to train the system on a video sequence for usage in tracking applications. Addition of a temporally consistent network would enable smooth detection and more optimal than per-frame detection.

REFERENCES:

Bruckner, Daniel. MLX-o-scope: a diagnostic visualization system for deep machine learning pipelines. No. UCB/EECS-2014-99.CALIFORNIA UNIV BERKELEY DEPT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCES, 2014.

K Saleh, Imad, Mehdi Ammi, and Samuel Szoniecky, eds. Challenges of the Internet of Things: Technique, Use, Ethics. John Wiley & Sons, 2018.

Petrov, Yordan. Improving object detection by exploiting semantic relations between objects.MS thesis. UniversitatPolitècnica de Catalunya, 2017.

Nikouei, SeyedYahya, et al. "Intelligent Surveillance as an Edge Network Service: from Harr-Cascade, SVM to a Lightweight CNN." arXiv preprint arXiv:1805.00331 (2018).

Thakkar, Kartikeya, et al. "Implementation and analysis of template matching for image registration on Devkit- 8500D." Optik-International Journal for Light and Electron Optics 130 (2017): 935-944..

Bradski, Gary, and Adrian Kaehler.Learning OpenCV: Computer vision with the OpenCV library." O'Reilly Media, Inc.", 2008.

Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).

Kong, Tao, et al. "Ron: Reverse connection with objectness prior networks for object detection." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).IEEE, 2017.

Liu, Wei, et al. "Ssd: Single shot multibox detector." European conference on computer vision.Springer, Cham, 2016.

Veiga, Francisco José Lopes. "Image Processing for Detection of Vehicles In Motion." (2018).

Huaizheng Zhang, Han Hu, GuanyuGao, Yonggang Wen, Kyle Guan, "Deepqoe: A Unified Framework for Learning to Predict Video QoE", Multimedia and Expo (ICME) 2018 IEEE International Conference on, pp. 1- 6, 2018.

Shijian Tang and Ye Yuan,"Object Detection based on Conventional Neural Network".

R. P. S. Manikandan, A. M. Kalpana, "A study on feature selection in big data", Computer Communication and Informatics (ICCCI) 2017 International Conference on, pp. 1-5, 2017 53 Warde-Farley, David. "Feedforward deep architectures for classification and synthesis." (2018).

Shilpisingh et al" An Analytic approach for 3D Shape descriptor for face recognition", International Journal of Electrical, Electronics, Computer Science & Engineering (IJEECSE), Special Issue - ICSCAAIT-2018| E-ISSN: 2348-2273 | P-ISSN: 2454-1222,pp-138-140.