



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Real Time Object Detection Using Deep Learning Techniques

Ankita Sontakke, Prof. Vijay Rakhade, Prof. Ashish Deharkar

Department of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India

Department of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India

Department of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India

ABSTRACT: Real time object detection is a vast, vibrant and complex area of computer vision. If there is a single object to be detected in an image, it is known as Image Localization and if there are multiple objects in an image, then it is Object Detection. This detects the semantic objects of a class in digital images and videos. The applications of real time object detection include tracking objects, video surveillance, pedestrian detection, people counting, self-driving cars, face detection, ball tracking in sports and many more. Convolution Neural Networks is a representative tool of Deep learning to detect objects using OpenCV (Opensource Computer Vision), which is a library of programming functions mainly aimed at real time computer vision.

KEYWORDS: Computer vision, Deep Learning, Convolution Neural Networks.

I. INTRODUCTION

The motive of object detection is to recognize and locate all known objects in a scene. Preferably in 3D space, recovering pose of objects in 3D is very important for robotic control systems. Imparting intelligence to machines and making robots more and more autonomous and independent has been a sustaining technological dream for the mankind. It is our dream to let the robots take on tedious, boring, or dangerous work so that we can commit our time to more creative tasks. Unfortunately, the intelligent part seems to be still lagging behind. In real life, to achieve this goal, besides hardware development, we need the software that can enable robot the intelligence to do the work and act independently. One of the crucial components regarding this is vision, apart from other types of intelligences such as learning and cognitive thinking. A robot cannot be too intelligent if it cannot see and adapt to a dynamic environment. The searching or recognition process in real time scenario is very difficult. So far, no effective solution has been found for this problem. Despite a lot of research in this area, the methods developed so far are not efficient, require long training time, are not suitable for real time application, and are not scalable to large number of classes. Object detection is relatively simpler if the machine is looking for detecting one particular object. However, recognizing all the objects inherently requires the skill to differentiate one object from the other, though they may be of same type. Such problem is very difficult for machines, if they do not know about the various possibilities of objects.

II. OBJECT DETECTION

Object Detection is the process of finding and recognizing real-world object instances such as car, bike, TV, flowers, and humans out of an images or videos. An object detection technique lets you understand the details of an image or a video as it allows for the recognition, localization, and detection of multiple objects within an image. It is usually utilized in applications like image retrieval, security, surveillance, and advanced driver assistance systems (ADAS). Object Detection is done through many ways:

Feature Based Object Detection

Viola Jones Object Detection

SVM Classifications with HOG Features

Deep Learning Object Detection

Object detection from a video in video surveillance applications is the major task these days. Object detection technique is used to identify required objects in video sequences and to cluster pixels of these objects. The detection of an object in video sequence plays a major role in several applications specifically as video surveillance applications. Object detection in a video stream can be done by processes like pre-processing, segmentation, foreground and background extraction, feature extraction. Humans can easily detect and identify objects present in an image. The

human visual system is fast and accurate and can perform complex tasks like identifying multiple objects with little conscious thought. With the availability of large amounts of data, faster GPUs, and better algorithms, we can now easily train computers to detect and classify multiple objects within an image with high accuracy..

III. DIGITAL IMAGE PROCESSING

Computerized picture preparing is a range portrayed by the requirement for broad test work to build up the practicality of proposed answers for a given issue. A critical trademark hidden the plan of picture preparing frameworks is the huge level of testing and experimentation that Typically is required before touching base at a satisfactory arrangement. This trademark infors that the capacity to plan approaches and rapidly model hopeful arrangements by and large assumes a noteworthy part in diminishing the cost and time required to land at a suitable framework execution.

Processing on image:

Processing on image can be of three types They are low-level, mid-level, high level.

Low-level Processing:

Preprocessing to remove noise.

Contrast enhancement.

Image sharpening.

Medium Level Processing :

Segmentation.

Edge detection

Object extraction.

High Level Processing:

Image analysis

Scene interpretation

RELATED TECHNOLOGY:

R-CNN

R-CNN is a progressive visual object detection system that combines bottom-up region proposals with rich options computed by a convolution neural network. R-CNN uses region proposal ways to initial generate potential bounding boxes in a picture and then run a classifier on these proposed boxes.

SINGLE SIZE MULTI BOX DETECTOR

SSD discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. At the time of prediction the network generates scores for the presence of each object category in each default box and generates adjustments to the box to better match the object shape. Additionally, the network combines predictions from multiple feature maps with different resolutions to naturally handle objects of various sizes.

ALEXNET

AlexNet is a convolutional neural Network used for classification which has 5 Convolutional layers, 3 Fully connected layers and 1 soft max layer with 1000 outputs for classification as his architecture.

YOLO

YOLO is real-time object detection. It applies one neural network to the complete image dividing the image into regions and predicts bounding boxes and possibilities for every region. Predicted probabilities are the basis on which these bounding boxes are weighted. A single neural network predicts bounding boxes and class possibilities directly from full pictures in one evaluation. Since the full detection pipeline is a single network, it can be optimized end-to-end directly on detection performance.

VGG

VGG network is another convolution neural network architecture used for image classification.

MOBILENETS

To build lightweight deep neural networks MobileNets are used. It is based on a streamlined architecture that uses depth-wise separable convolutions. MobileNet uses 3×3 depth-wise separable convolutions that uses between 8 times less computation than standard convolution at solely a little reduction accuracy. Applications and use cases including object detection, fine grain classification, face attributes and large scale-localization.

TENSOR FLOW

Tensor flow is an open source software library for high performance numerical computation. It allows simple deployment of computation across a range of platforms (CPUs, GPUs, TPUs) due to its versatile design also from desktops to clusters of servers to mobile and edge devices. Tensor flow was designed and developed by researchers and engineers from the Google Brain team at intervals Google's AI organization, it comes with robust support for machine learning and deep learning and the versatile numerical computation core is used across several alternative scientific domains.

OPEN CV

OpenCV stands for Open supply pc Vision Library is associate open supply pc vision and machine learning software system library. The purpose of creation of OpenCV was to produce a standard infrastructure for computer vision applications and to accelerate the utilization of machine perception within the business product [6]. It becomes very easy for businesses to utilize and modify the code with OpenCV as it is a BSD-licensed product. It is a rich wholesome library as it contains 2500 optimized algorithms, which also includes a comprehensive set of both classic and progressive computer vision and machine learning algorithms. These algorithms is used for various functions such as discover and acknowledging faces. Identify objects classify human actions. In videos, track camera movements, track moving objects. Extract 3D models of objects, manufacture 3D purpose clouds from stereo cameras, sew pictures along to provide a high-resolution image of a complete scene, find similar pictures from a picture information, remove red eyes from images that are clicked with the flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality

IV. APPLICATION OF OBJECT DETECTION

The major applications of Object Detection are:

FACIAL RECOGNITION

“Deep Face” is a deep learning facial recognition system developed to identify human faces in a digital image. Designed and developed by a group of researchers in Facebook. Google also has its own facial recognition system in Google Photos, which automatically separates all the photos according to the person in the image. There are various components involved in Facial Recognition or authors could say it focuses on various aspects like the eyes, nose, mouth and the eyebrows for recognizing a faces

PEOPLE COUNTING

People counting is also a part of object detection which can be used for various purposes like finding person or a criminal; it is used for analysing store performance or statistics of crowd during festivals. This process is considered a difficult one as people move out of the frame quickly.

INDUSTRIAL QUALITY CHECK

Object detection also plays an important role in industrial processes to identify or recognize products. Finding a particular object through visual examination could be a basic task that's involved in multiple industrial processes like sorting, inventory management, machining, quality management, packaging and so on. Inventory management can be terribly tough as things are hard to trace in real time. Automatic object counting and localization permits improving inventory accuracy.

SELF DRIVING CARS

Self-driving is the future most promising technology to be used, but the working behind can be very complex as it combines a variety of techniques to perceive their surroundings, including radar, laser light, GPS, odometer, and computer vision. Advanced control systems interpret sensory info to allow navigation methods to work, as well as obstacles and it. This is a big step towards Driverless cars as it happens at very fast speed.

SECURITY

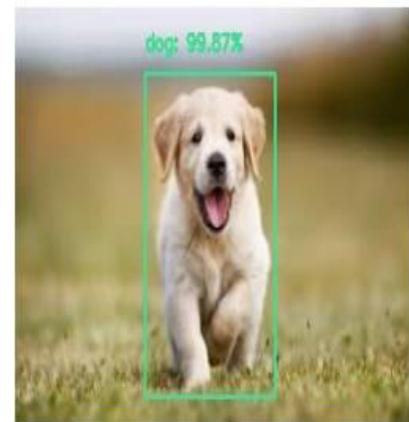
Object Detection plays a vital role in the field of Security; it takes part in major fields such as face ID of Apple or the retina scan used in all the sci-fi movies. Government also widely use this application to access the security feed and match it with their existing database to find any criminals or to detecting objects like car number involved in criminal activities. The applications are limitless.

V. RESULTS

Input



Output





VI. CONCLUSION

Deep-learning based object detection has been a search hotspot in recent years. This project starts on generic object detection pipelines which give base architectures for other related tasks. With the assistance of this the 3 other common tasks, namely object detection, face detection and pedestrian detection, are often accomplished. Authors accomplished this by combing 2 things: Object detection with deep learning and OpenCV and Efficient, threaded video streams with OpenCV. The camera sensor noise and lightening condition can change the result because it can create problem in recognizing the objects. generally, this whole process requires GPU's rather than CPU's. But we've done using CPU's and executes in much less time, making it efficient. Object Detection algorithms act as a mixture of both image classification and object localization. It takes the given image as input and produces the output having the bounding boxes adequate to the amount of objects present within the image with the category label attached to every bounding box at the highest. It projects the scenario of the bounding box up the shape of position, height and width.

VII. FUTURE SCOPE

As Object detection and recognition in todays world can be considered as one of the most challenging, complex and most important task in computer vision Field. As we know that this project is been developed with the underlying purpose of real-time object in pictures, videos captured streaming cameras. or web cams.

- It can be improve and innovate in the future by anybody without worrying about complexity.
- Future enhancements can be focused by implementing the project on the system having GPU for faster results and better accuracy.
- Like, for small object detection which is done by MS COCO in some face detection application and task. For improvement of localization of small objects under partial barrier. So that we will improve the network architecture with some modifications.
- By using that we can reduce the dependency of data network.
- For achieving efficient recognition of small objects with better accuracy.

So it is finally concluded that for enhance the accuracy and performance by using pre processing techniques like edge detection and increase image augmentation and contrast so that we get better results in output.

REFERENCES

- [1] Bruckner, Daniel. MI-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.
- [2] 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.Universitat Politècnica de Catalunya, 2017.
- [3] 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).

- [4]Thakar, Kartikey, 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..
- [5] Bradski, Gary, and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc., 2008.
- [6] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).
- [7]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.
- [8] Liu, Wei, et al. "Ssd: Single shot multibox detector." *European conference on computer vision*.Springer, Cham, 2016.
- [9]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 Conferenceon*, pp. 1- 6, 2018.
- [10] Lowlesh Yadav and Asha Ambhaikar, "IOHT based Tele-Healthcare Support System for Feasibility and performance analysis," *Journal of Electrical Systems*, vol. 20, no. 3s, pp. 844–850, Apr. 2024, doi: 10.52783/jes.1382.
- [11] L. Yadav and A. Ambhaikar, "Feasibility and Deployment Challenges of Data Analysis in Tele-Healthcare System," *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*, Raipur, India, 2023, pp. 1-5, doi: 10.1109/ICAIIHI57871.2023.10489389.
- [12] L. Yadav and A. Ambhaikar, "Approach Towards Development of Portable Multi-Model Tele-Healthcare System," *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*, Raipur, India, 2023, pp. 1-6, doi: 10.1109/ICAIIHI57871.2023.10489468.
- [13] Lowlesh Yadav and Asha Ambhaikar, *Exploring Portable Multi-Modal Telehealth Solutions: A Development Approach*. *International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC)*, vol. 11, no. 10, pp. 873–879, Mar. 2024.11(10), 873–879, DOI: 10.13140/RG.2.2.15400.99846.
- [14] Lowlesh Yadav, *Predictive Acknowledgement using TRE System to reduce cost and Bandwidth*, March 2019. *International Journal of Research in Electronics and Computer Engineering (IJRECE)*, VOL. 7 ISSUE 1 (JANUARY-MARCH 2019) ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE).



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details