Object Detection using PYTHON Programming

Dr.Vithya Ganesan[1], Dr.V.Anjana Devi[2], Dr.Pachipala Yellamma[3]

[1-3]Professor, Department of Computer Science Engineering, Koneru Lakshmiah Education Foundation

Vaddeswaram, Andra Pradesh, India.

Email id: vithyaganesan@kluniversity.in, 170030338@kluniversity.in

Abstract

The capacity for cameras and computational frameworks to screen a given region has been set up and worked upon by the specialist for quite a long while. The capacity of the human mind to just classify and comprehend what an object is upon a basic look is something that has interested numerous scientists for quite a while. Indeed, even with the progressions in the areas of Artificial insight also ML is as yet a troublesome errand for researchers to make an object following and recognition framework that can work productively in rapidly. We have had the option to make models in the realm of ML which have been prepared thoroughly with a huge number of pictures inside datasets to perceive and comprehend a few objects however we have still neglected to appropriately set up one equivalent to that of the human mind. Although, the downside of computational elements is their capacity to share feelings we have likewise had the option to comprehend the human brain has undeniably more computational force then one can envision. In our task, we will be pursuing making a calculation that will actually want to perceive and portray the entirety of the pictures inside a given edge. To achieve this, we will be building up a module utilizing Python and a webcam that catches the pictures in the type of casings and perceive objects. We will make a calculation that will contrast the object inside the casing with that of a dataset set up inside Amazon AWS to perceive and comprehend what the object is. Through this paper, we will discuss the sorts of programming used to execute to make this conceivable alongside our discoveries alongside the future extent of this innovation.

Keywords: *Object detection, Object recognition, Open CV, Python, Machine learning.*

I.   Introduction

The ability for the human mind to immediately recognize an object even with the most minute details is something which is astonishing. Scientist have tried to understand the neural networks in a human mind to gain an understanding upon how this is accomplished but this is one of those phenomena which cannot be explained or understood in a technical aspect [1].

Over the past decade we have been introduced within various new forms of technologies such as AI as well as ML which have allowed us to create computational entities which are capable of performing human tasks and jobs in an efficient manner. However, they have their own limitations as they are restricted with their computational power to only understand what they have been trained to perform. Due to this they are unable to express human emotions or traits within them. Having taken this aside these technologies have enabled us to create various different forms of computational models which are able to compete on par with human intelligence. For example, in recent times researchers have been able to create a chess AI bot capable of even defeating the best chess player in the world. Even with the amount of progress that has been shown in these industries they still fail to

achieve object recognition on par with that of the human mind. Modern recognition models are able to look into various objects and have them properly identified however their ability to do so in a fraction of a second as that of the human mind is still lacking. Furthermore, the model will require lots of details and maybe even the entire object to detect an recognize the object whereas the human mind can identify an object with even the most minute details or portion of the object. Having worked upon this for years scientists have still been unable to develop and create an object recognition model which can compete with humans however there have been several different models which have been created with pleasing accuracies above 85%.

Although modern recognition algorithms are able to properly track the objects within the frame and have them properly identified it is often inaccurate and can take large amounts of time to create such models. In order to train our model to recognize a single object it will take large amounts of time as well as datasets. Furthermore, our model will be unable to recognize an object if it is oriented in a different manner or missing any parts of itself. Most models which have been created in recent times by various different people around the world often take several hours to train as well as large datasets to accomplish. Even with all of these provided these models often tend to have accuracies lower than 90%. This is why we would like to take these as the challenges to overcome within the model which we will be creating.

Instead of spending large amounts of time to manually train our model we will be using the datasets established within in Amazon AWS cloud to compare and recognize images. This dataset which has been created and established any AWS has various numbers of objects within it so that our algorithm will not bump into unidentifiable objects. As we have already discussed the major advantage to following this method is its ability to recognize objects with absolutely no training, we were further able to receive results which have surprised us. The results for our algorithm will be discussed in the later parts of our paper however let's take a look into the types of hardware which we had to implement and use to create this prototype.

Through this paper you will be able to look into how this has been established along with the results we have been able to obtain. We will further talk about the future scope of our project and how it can be developed even further in the future in a much larger scale.

II. Literature Review

Hasan Hashib; Md. Leon et al This paper talks about the execution of a practical, keen security framework that defeats the downsides of regular surveillance cameras by using an AI and Viola-Jones calculation under picture preparing writing to distinguish intruders and numerous object detection continuously. The paper presents the plan and execution subtleties of the savvy object detection based security framework in two distinctive processing climate, MATLAB and Python separately utilizing Raspberry Pi 3 B single-board PC. The security framework is fit for cautioning the security admin through email by means of the web while initiating an alert locally. [2].

Schalk Wilhelm Pienaar; Reza Malekian et al., explained about the motivation behind this paper is to talk about an approach to prepare a model that can limit and catch the conditions of underground diggers utilizing a Single Shot Detector (SSD) model, prepared explicitly to make a differentiation between a harmed and a non-harmed excavator (resting versus holding up). Tensorflow is utilized for the deliberation layer of executing the AI calculation, and despite the fact

that it utilizes Python to manage hubs and tensors, the real calculations run on C++ libraries, giving a decent harmony among execution and speed of improvement. The paper further talks about assessment techniques for determining the exactness of the AI progress. For future work, data fusion is acquainted all together with improve the exactness of the recognized action/condition of individuals in a mining climate. [3].

J. Talukdar; S. Gupta et al., investigated various procedures to create manufactured datasets and consequently improve them to accomplish better object detection exactness (mAP) when prepared with best in class profound neural networks, zeroing in on detection of pressed food items in a cooler. They created novel procedures like dynamic stacking, pseudo arbitrary arrangement, variable object present, distractor commotion and so on which help in differentiating the manufactured data as well as help in improving the general object detection mAP by over 40%. The manufactured pictures, produced utilizing Blender-Python API, are bunched in an assortment of arrangements to oblige the variety of genuine scenes. These datasets are then used to prepare TensorFlow executions of best in class profound neural networks like Faster-RCNN, R-FCN, and SSD and their exhibition is tried on genuine scenes. The object detection execution of different profound CNN designs is additionally contemplated, with Faster-RCNN ending up being the most appropriate decision, accomplishing the most elevated mAP of 70.67 [4].

Hasan UCUZAL et al., explained that Object recognition is a PC vision strategy for distinguishing objects in pictures or recordings. Object recognition is a significant yield of profound learning and AI calculations. For this reason, open source, free and man-made consciousness based "Object Recognition Software" has been created to perform object recognition activity without any problem. In making this web-based programming, Darkflow and Tensorflow libraries are utilized which depend on profound learning based Python programming language and permit the plan of intuitive web based applications. While performing object recognition examination in the created programming, CNN (Convolutional Neural Networks) numerous convolution layers are uncovered covered up and helpful highlights got by different figuring strategies. With CNN, objects are characterized, objects are recognized, and objects are controlled by picture division. A pre-prepared model from COCO, a huge scope object detection, dividing and picture dataset, is utilized to perceive how the web-based programming work and to assess the investigation yields. Object recognition investigation is applied to ten pictures from this data set. As indicated by the object recognition examination aftereffects of the ten pictures, the determined precision rates is inspected and it is tracked down that this web based programming which is created as open source and free access gives fruitful assessments in object recognition. To perceive how the web-based programming functions and to assess the investigation yields, a pre-prepared model was utilized from COCO (Common Objects in Context) which is an enormous scope object detection, parceling and picture dataset. Object recognition examination was applied to ten pictures from this data set. At the point when the exactness proportion of the ten pictures determined by the object recognition examination result is inspected, it is resolved that this web based programming which is created as open source and free access gives fruitful forecasts in object recognition. The created programming is new easy to use web-based programming that can without much of a stretch recognize objects in pictures and prejudicial from one another objects. In the accompanying investigations, to build the analytic exactness of the objects in the pictures, it is proposed that the programming projects that utilizes

further neural networks ought to be created and the important foundation to recognize the deformities in the clinical pictures can be created [5].

Zhong-Qiu et al., gave a survey on profound learning based object detection structures. Our audit starts with a concise presentation on the historical backdrop of profound learning and its delegate instrument, specifically Convolutional Neural Network (CNN). At that point we center around commonplace nonexclusive object detection structures alongside certain alterations and valuable stunts to improve detection execution further. As unmistakable explicit detection undertakings show various qualities, we likewise momentarily study a few explicit undertakings, including remarkable object detection, face detection and person on foot detection. Test examinations are likewise given to think about different strategies and draw a few significant ends. At long last, a few promising bearings and undertakings are given to fill in as rules to future work in both object detection and significant neural organization based learning frameworks. [6].

Bernardo Augusto Godinho de Oliveira et al., explained the characteristic capacity of people to quickly recognize, separate, and arrange objects permits us to settle on brisk choices with respect to what we see. A few apparatuses can utilize quick and lightweight robotized object detection for pictures or recordings. All through the most recent five years, the innovation business has continually presented computational and equipment arrangements, like gadgets with great handling and capacity abilities. Be that as it may, object detection strategies ordinarily require either high preparing force or huge stockpiling accessibility, making it hard for asset obliged gadgets to play out the detection progressively without an association with an incredible worker. The model introduced in this paper requires just 95 megabytes of capacity and took 113 ms in normal for every picture running on a PC CPU, making it reasonable for independent gadgets that can be utilized in a hurry. [7].

Luis Barba-Guamán et al., disclosed some hypothetical ideas to recognize objects by methods for their shading (thresholding), this procedure was carried out in the advancement of a game program. Moreover, the thresholding range for the red, yellow and green tones was found to accomplish a superior methodology in the object detection. This task utilized the python programming language, Pygame graphical interface libraries and the OpenCV library free open source about counterfeit vision. [8].

Gerardo Asael López-Alfaro et al., presented the plan of a portable robot which utilizes an Internet of Things design to be controlled distantly. This robot incorporates a camera which is utilized to get scene pictures from the portable robot. A web worker permits to store the orders to control the developments of the portable robot in a MySQL database just as the transmission of orders between the worker and the versatile robot; by utilizing web pages with PHP and an Android APP. Their electronic framework depends on WiFi Module ESP8266 which offers benefits to join open programming designs. The scene pictures are prepared in a PC utilizing python and AI libraries to distinguish fascinating objects on the pictures. The AI calculations perceive on the scene pictures two kind of objects classes: agave plants and stones. We utilize a convolutional neural organization for the object detection known as You Only Look Once (YOLO) which accomplishes 81.2% of precision for perceiving the two classes. [9].

Chandan G et al., explained, Deep learning has acquired an enormous impact on how the world is adjusting to Artificial Intelligence since recent years. A portion of the mainstream object detection calculations are Region-based Convolutional Neural Networks (RCNN), Faster-RCNN, Single Shot Detector (SSD) and You Only Look Once (YOLO). Among these, Faster-RCNN and SSD have

better precision, while YOLO performs better when speed is given inclination over exactness. Deep learning joins SSD and Mobile Nets to perform effective execution of detection and following. This calculation performs effective object detection while not settling on the presentation [10].

III.   Proposed Methodology

A. Software Requirements
➢ Apache NetBeans
➢ Python plugins
➢ JDK 1.8

B. Technology Requirements
The Python libraries used in this project include-

➢ Pyttsx3

This is a text-to-speech conversion in-built python library that can be used to convert text data into speech.

➢ Csv

Csv (Comma separated values) in an inbuilt python library used for file reading and writing. The csv module is used to read and write data in tabular form to csv format [11].

➢ Cv2

Cv2 is an Open-cv related in-built python library that is used to solver computer vision related problems. It fundamentally centers around image processing, video catch and examination including highlights like face detection and object detection [12].

➢ Boto3

Boto3 is the Amazon Web Services (AWS) Software Development Kit (SDK) for Python, which helps Python developers to write software that uses services such as Amazon S3 and Amazon EC2.

C. Implementation/Flow of the project
The project mainly concentrates on capturing images continuously till the user stops the model. The objects captured during the session are then analyzed and are generated with identification and confidence interval.

The flow of the Object detection model goes like-

Step 1: Importing the required python libraries

Libraries used include-

▪ Csv
▪ Boto3
▪ Cv2
▪ Pyttsx3

Step 2: Opening cam and enabling video capture

In this module, we use OpenCV libraries to enable camera and record series of images in the form of frames

Step 3: Accessing AWS dataset

Using boto3 and pyttsx3 libraries, we now get the capacity to access the dataset from AWS or we can simply go with step 2 to continue with the dataset.

The dataset includes access_key_id and secret_access_key for every image. Meanwhile, the model starts reading the images from the dataset and returns frames of images.

The model displays "failed to grab frame" whenever the model does not return image frame.

Step4: Object recognition

From the captured frames, by evaluating the model, the program generates labels from the responses recorded and compares it from the database. Then, it generates the name of the object detected.

Later on, the confidence level of the identified object is generated. The minimum confidence interval of the developed model is 95. A confidence model is a method of adding confidence stretch data to a predictive model. Genuinely, for a given forecast, a confidence model furnishes a stretch with upper and lower limits, inside which it is sure, up to a specific level, that the real worth happens. During predictive scoring, this proportion of confidence gives extra data about the accuracy of the expectation. Confidence models work diversely relying upon the kind of information being anticipated. It returns the confidence stretch for a population mean, utilizing a typical appropriation.

The confidence stretch is a scope of qualities. Your example mean, x, is at the focal point of this reach and the reach is x ± CONFIDENCE. For instance, if x is the example mean of conveyance times for items requested through the mail, x ± CONFIDENCE is a scope of population means [13].

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}} \tag{1}$$

Where,

CI-confidence interval
$\bar{x}$- sample mean
z- confidence level value
s- sample standard deviation
n-sample size

In this project, we evaluated the confidence interval of the model.

The Object detection model developed in this project detects different kinds of objects. The figure below shows detection of a Person and face of a person with confidence level of 97.91
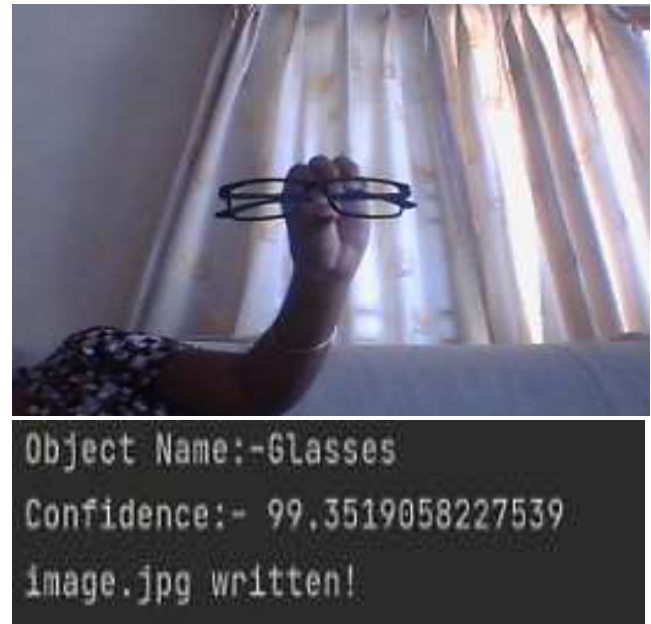
IV.   Objects and Outputs

Similarly, here is another object (glasses) that has been detected by the model with a confidence level of 99.35
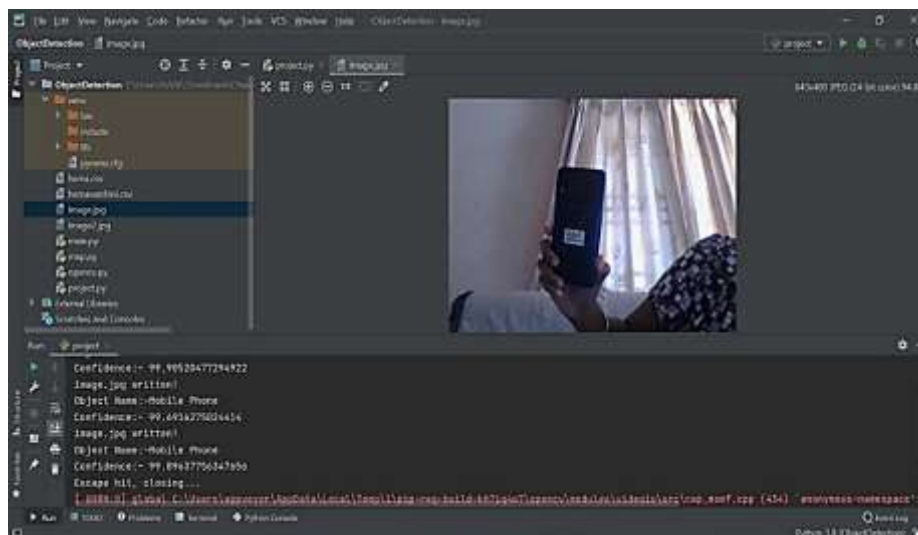

Fig 1. Image of a person detected with its confidence interval.

Fig 2. Image of glasses detected with its confidence interval.

Here goes another object (mobile phone) that has been detected with a confidence level of 99.89%.

Fig 3. Image of a mobile phone detected with its confidence interval

V.   Conclusion

Through our undertaking, we have had the option to figure out how object following, just as object detection, works inside this present reality. We have had the option to effectively make not just a calculation equipped for following and distinguishing objects with a confidence level above

95%. Although the results of our model have astounded us, we accept that the degree whereupon it very well may be created is somewhat enormous. One of the significant achievements inside our ventures is the capacity of our calculations to work productively with high exactness with no preparation by any stretch of the imagination. Moreover, it can follow as mark objects with high productivity. We accept that later on we might have the option to additionally build up this task in different various strategies. Later on, we might have the option to utilize a similar calculation to make a model fit for checking objects in a lot bigger casing. Moreover, with more computational force we may even have the option to distinguish just as mark many various objects inside a given edge. Object recognition may never be able to rival the capacities of human behaviour yet we accept that we may contrast them and each other and create upon these to reach much bigger stages.

References

[1] Prakhar Ganesh. Object Detection: Simplified(2019). Retrieved from https://towardsdatascience.com/object-detection-simplified-e07aa3830954.

[2] Hashib, Hasan, Md Leon, and Ahmed Mortuza Salaque. "Object Detection Based Security System Using Machine learning algorthim and Raspberry Pi." *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*. IEEE, 2019.

[3] Pienaar, Schalk Wilhelm, and Reza Malekian. "Human activity recognition using visual object detection." *2019 IEEE 2nd Wireless Africa Conference (WAC)*. IEEE, 2019.

[4] Talukdar, Jonti, et al. "Transfer learning for object detection using state-of-the-art deep neural networks." *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 2018.

[5] UCUZAL, Hasan, et al. "A web-based application for identifying objects in images: object recognition software." *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2019.

[6] Zhao, Zhong-Qiu, et al. "Object detection with deep learning: A review." *IEEE transactions on neural networks and learning systems* 30.11 (2019): 3212-3232.

[7] de Oliveira, Bernardo Augusto Godinho, Flávia Magalhães Freitas Ferreira, and Carlos Augusto Paiva da Silva Martins. "Fast and lightweight object detection network: Detection and recognition on resource constrained devices." *IEEE Access* 6 (2018): 8714-8724.

[8] Barba-Guamán, Luis, Carlos Calderon-Cordova, and Pablo Alejandro Quezada-Sarmiento. "Detection of moving objects through color thresholding." *2017 12th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2017.

[9] López-Alfaro, Gerardo Asael, et al. "Mobile robot for object detection using an IoT system." *2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*. Vol. 4. IEEE, 2020.

[10] Chandan, G., Ayush Jain, and Harsh Jain. "Real time object detection and tracking using Deep Learning and OpenCV." *2018 International Conference on inventive research in computing applications (ICIRCA)*. IEEE, 2018.

[11] CSV File Reading and Writing. Retrieved from https://docs.python.org/3/library/csv.html

[12] Open CV Tutorial. Retrieved from https://www.tutorialspoint.com/opencv/index.html

[13] Confidence Models. Retrieved from http://support.ptc.com/help/thingworx_hc/thingworx_analytics_8/index.html#page/analytics/confidence_models.html