

Object and its dimension detection in real time

Madhavi Karanam¹, Varun Kumar Kamani^{1,*}, Vikas Kuchana¹, Gopal Krishna Reddy Koppula¹, and Gautham Gongada¹.

¹Department of CSE, Gokaraju Rangaraju Institute of Engineering and Technology, India

Abstract. Object and its dimension detection from images and videos can be very helpful for everyday use. This paper discusses the use of the system to detect an object in real time and provide its dimensions upon demand. The object dimension measurement and detection are some of the important topics of computer vision which helps in automating the manual tasks. Human beings are capable of recognizing and spotting objects in images and videos, but computers lack that ability with out prior training. To train the computer, we must use machine learning, computer vision, and object detection algorithms. This project provides the way to detect and measure an object's dimension in real time from a webcam. To estimate the object's dimension in real time, we have utilized the OpenCV and NumPy libraries. Computer vision provides support to computers to observe and understand. Computer vision helps the computer in understanding a 3D surrounding from a 2D image and trains the computer to perform different functions. It also helps in Human Computer Interaction effectively because it is able to differentiate the objects with surroundings and provide us with the key information.

1 Introduction

The study of how computers perceive and understand digital images and videos is called computer vision. Computer vision embodies all of the activities provided by the biological visual system, which includes "seeing" or interpreting visual stimuli, understanding what is being observed, and distilling complex concepts into a format that will be used in a variety of relevant and useful processes. This multidisciplinary field uses sensors, computers, and machine learning techniques to imitate and automate these components of human visual systems. Object Detection comes under Computer vision technology. Computer vision provides the machine to have the ability to observe, understand, locate certain things in the images or videos. Object Detection will also provide help to track the objects and we can categorize the objects we have discovered and find examples of those items in the real time scenario. There are several applications for object identification and tracking in computer vision research, including traffic detection, vehicle navigation, and interpersonal connections. object detection in digital photos and videos refers to the process of finding examples of semantic items of a specific class utilizing computer equipment connected to computer vision and image processing

* Corresponding author : kamani.varunkumar@gmail.com

Three of the numerous applications for object detection include face detection, face recognition, and video objects. The object detection system detects whether things are present or absent in particular scenes and from particular camera angles. The diverse object detection domains are based on varied goals and are categorized using concrete and abstract terms. The various models that are either expressly or implicitly employed for object detection. The components may or may not vary depending on the different techniques, selecting an item based on a hypothesis and selecting an item based on matching. The most effective processing technique is object detection. Identifying objects in a situation is the process of object detection in practical settings.

This paper aim is to propose an approach that can detect an object and provide its dimensions. The measurements can be utilized in various applications, such as automatic parking solutions, self-driving automobiles, and e-commerce platforms (such as Myntra, Ajio, Amazon, Flipkart, etc.) to determine the size of the products to best suit the needs of customers. The discussion of related work has been done in Section 2. Section 3 describes the proposed approach, Section 4 presents the proposed approach's results, and Section 5 concludes the paper.

2 Related work

In this literature survey, we review various papers that explore different approaches for object detection. An approach for real-time object detection and tracking using deep learning with the help of OpenCV used a pre-trained deep neural network, namely You Only Look Once (YOLO), which is designed to detect objects in real-time. The proposed method achieved an accuracy of 92.6% on the PASCAL VOC 2012 data-set [1]. A review of research on object detection based on deep learning surveyed recent advancements in object detection, including Faster R-CNN, YOLO, SSD, and RetinaNet, it also discusses various datasets used for object detection and the evaluation metrics used to measure the performance of object detection algorithms [2]. The paper studied object detection and proposed an approach that used a combination of image processing techniques and machine learning algorithms to detect and classify objects. Various methods are researched in this, and calculated accuracy of the methods. The final proposed method achieved an accuracy of 87.5% on a custom datasets [3]. A comparative study of object detection algorithms, including Haar cascades, Histogram of Oriented Gradients (HOG), and deep learning-based methods such as YOLO, Faster R-CNN, and SSD. The study compared the accuracy, speed, and computational complexity of these algorithms [4]. Another paper proposed an efficient object detection algorithm that achieved state-of-the-art results on the COCO data-set. The authors used a compound scaling method to balance accuracy and efficiency and proposed a new scaling rule to determine the depth and width of the network [5].

A study of object detection methods and their applications on digital images. The authors discussed various object detection algorithms, including traditional computer vision methods such as Haar cascades and HOG, and deep learning-based methods such as YOLO and Faster R-CNN [6]. Another study proposes a system for detecting and tracking objects using UAVs. The system combines several image processing techniques, including feature extraction, object segmentation, and optical flow analysis, to identify objects in real-time. The system's performance is evaluated using a data-set of UAV images, and the results show that it can achieve high accuracy in detecting and tracking objects, even in challenging environments [7]. A method for texture classification that can be used for object detection. The method uses local binary patterns to extract features from images at multiple scales and orientations, which are then used to train a classifier. The paper presents experimental results that show the effectiveness of the proposed method in detecting textures and objects in images [8]. Object detection algorithm called faster R-CNN, that can

achieve high accuracy while maintaining real-time performance. The algorithm uses a convolutional neural network (CNN) to extract features from images, which are then used to generate object proposals. The object proposals are refined using a region-based CNN, which produces a final set of object detection. The paper presents experimental results that show the proposed algorithm outperforms previous state-of-the-art methods in terms of accuracy and speed [9]. Another algorithm called, You Only Look Once (YOLO) Unified, Real-Time Object Detection, which presents an object detection algorithm that can detect objects in real-time with high accuracy. The algorithm uses a single neural network to predict object classes and bounding boxes directly from input images, eliminating the need for region proposals. The paper presents experimental results that demonstrate the effectiveness of the proposed algorithm in detecting objects in various datasets [10].

OpenCV library, which is a free, open-source computer vision library designed for real-time applications. The paper outlines the motivation for developing OpenCV, the key features of the library, and its architecture. The author explains that computer vision has traditionally been a complex and challenging field due to the diversity of image and video processing applications. OpenCV was designed to provide a common platform for computer vision research and development. The library is intended for use in real-time applications such as robotics, surveillance, and image and video processing. The paper describes the key features of OpenCV, including its image and video processing functions, feature detection and extraction algorithms, and machine learning capabilities [11].

After observing the various research studies, we decided that the computer vision is the most suitable technology for detecting the objects in our project. In conclusion, these papers demonstrate the potential of advanced techniques and algorithms for object dimension detection. However, there is still room for improvement, as this is a very modern field of study. The object detection detection is also a contributor of advancement of robotics and AI.

3 Proposed approach

In this paper, we propose an approach that uses the video or image of the surroundings captured by the computer's webcam or external camera as an input to identify items, measure their dimensions, and identify them. With a stand, we are able to keep the camera at a specific distance while also adjusting the camera's height, width, and depth. With the use of this system, we are able to identify many objects at once, provide their dimensional measurements, and obtain other information such as the object's area of occupancy. The system is created using a python program which makes use of python libraries for queue, math, numpy, and computer vision (opencv-cv2). This system is composed of three modules (a, b and c) that act in accordance with various system-related tasks and are discussed below sequentially.

Module a - the first module is used for obtaining the video frames input, in this module we configure our environment by setting the camera width, height and also we set the frame rate of the video input.

Module b - the second module is used to detect object by capturing the objects boundaries. There are multiple functions performed in this module. Initially the input frame to gray scale to get better understanding of the details in input. Then we apply Gaussian blurring this methodology substitutes a Gaussian kernel for a box filter. The cv.GaussianBlur() function has completed its work. The kernel's width and height should be positive and odd, accordingly. Furthermore, we must supply sigmaX and sigmaY, which stand for the standard deviations in the X and Y axes, respectively. sigmaY is assumed to be the same as sigmaX if only sigmaX is supplied. If either or both are provided as zeros, values are chosen based on the kernel size. In order to exclude Gaussian noise from an

image, Gaussian blurring is highly effective. Then image thresholding `cv.threshold()` is applied, the thresholding is achieved through using procedure `cv.threshold`. The provided image, which must be in grayscale, is the first argument. The threshold value, which is used to classify the pixel values, is the second input. The following input determines the highest value that will be given to pixel values that are higher than the threshold. The function's final parameter, which is obtained using OpenCV, supports multiple types of thresholding. Then contours are detected three choices exist when employing the `cv.findContours()` function: source picture, contour retrieval mode, and contour approximation method. And the hierarchy and contours are produced. List of every contour in the image is identified as contours. The (x,y) coordinates of an object's border points are arrayed for each contour.

Module c - the third module is used to find the object dimensions by using the contour data we find the various lengths of the object from the math library we use `hypot` method which is used to calculate euclidean distance, there by allowing us to calculate length, breadth, and also also area of the object.

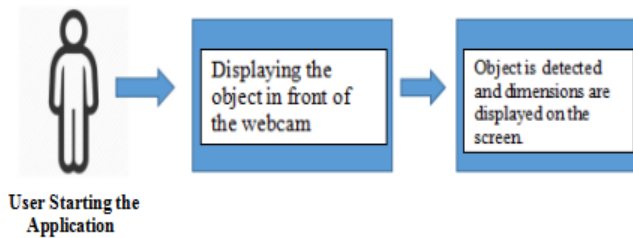


Fig. 1. Block diagram of proposed approach

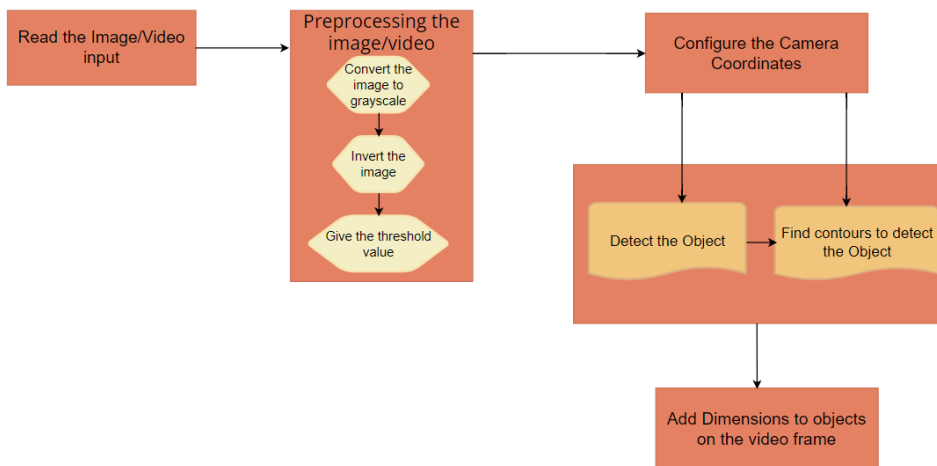


Fig. 2. Block diagram of working procedure

4 Results

Here we present sample outputs from our proposed approach involving various objects, and in table 2 we determine the accuracy using the results of the object dimensions. In Fig. 3 it

is showing mobile phone is detected and screen is displaying the dimensions and it's area. In Fig. 4 it is showing pen is detected and screen is displaying the dimensions of pen and it's area.

Table 2 is showing the results for various objects we tested. In this table, we recorded the measurements of different objects and compared with the actual dimensions of the objects and then calculated the percentage error to get the accuracy of the measurement.

$$\text{Percentage Error} = \frac{|OA - AA|}{|AA|}$$

OA=Observed Area

AA=Actual Area

$$\text{Accuracy}(\%) = 100 - \text{Percentage Error}$$



Fig. 3. Sample output 1

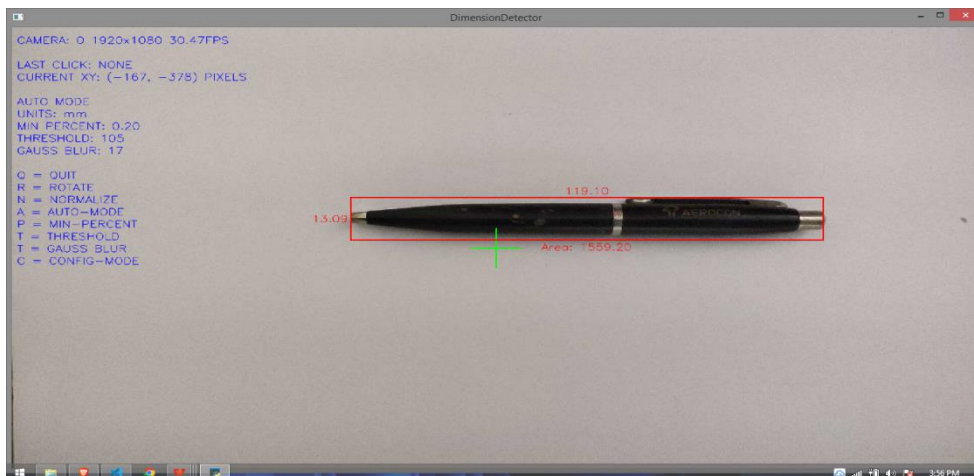


Fig. 4. Sample output 2

Table 1. Results of Proposed system being tested on various objects

Object	Actual Length (mm)	Actual Breadth (mm)	Actual Area (mm*mm)	Observed Length (mm)	Observed Breadth (mm)	Observed Area (mm*mm)	Accuracy (%)
Mobile	155	72	11160	153.29	70.29	10774.47	96.54
Pen	120	13	1560	119.10	13.09	1559.19	99.95
Book	275	164	45100	276.01	163.5	45127.63	99.94
Box	302	305	92110	294.50	310.1	91324.45	99.14
Wallet	91	73	6643	92.74	74.69	6926.75	95.72
Pencil Box	140	45	6300	141.1	45.0	6349.5	99.21
Adapter	34	47	1598	34.99	48.46	1695.61	93.89

5 Conclusion

In this paper, we have used computer vision because it helps computers to process, analyze, and understand the digital videos. The proposed work shows the application of computer vision that exhibits a significant level of accuracy for object detection and dimension measurement. The suggested work is used in getting the dimensions of the objects with an accuracy of over 95% most of the times, these dimensions are used in calculation of the area occupied by the object. Our proposed work can find various applications, including every day use, e-commerce applications, Industrial use and AI Automobiles , e.t.c.

References

1. G. Chandan, Mohana and A. Jain, Real-Time Object Detection and Tracking by Using OpenCV and Deep Learning, ICIRCA, (2018).
2. W. Wang, Z. Li, J. Deng, H. Yao, X. Xuan, Z. Wang, A review of research on an object detection based on deep learning, AINIT, (2020).
3. S.Manjula, Dr. K. Lakshmi, A study on object detection, IJPTFI,ISSN: 0975-766X, (2016).
4. U. Binay , N. Yadav , Comparative Study of Object Detection Algorithms, IRJET-V4I1103, Volume 4, Issue 7, (2017).
5. Brain Team, V. Quoc Le, Google Research, R. Pang, M. Tan, EfficientDet: Scalable and Efficient Object Detection, CPR, (2020).
6. Sunil, Gangandeep, Study of object detection methods and applications on digital images, IJSDR1905088, Volume 4, Issue 5, (2019).
7. K. Shreyamsh, UAV: Application of object detection and tracking techniques for unmanned aerial vehicles, Texas A & M University, (2015).
8. M. Pietikinen, T. Ojala, T. Maenpaa, Multi-resolution gray scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 24, Issue 7, 971 - 987, (2002).
9. R. Girshick, Fast R-CNN, In Proceedings of the IEEE international conference on computer vision pp. 1440-1448, (2015).
10. R. Girshick, , S. Divvala, J. Redmon and A. Farhadi, (YOLO) You only look once: Unified, real-time object detection, arXiv preprint arXiv:1506.02640, (2015).
11. G. Bradski, OpenCV: A general purpose computer vision library, Computer Vision, (2008).