Evolution in Electrical and Electronic Engineering Vol. 2 No. 2 (2021) 413-416

EEEE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

Real-Time Object Detection and Localization System with Mild Artificial Intelligence

Bong Yong Kang[1], **Xavier Ngu Toh Ik**[1]*

[1]Department of Electronic Engineering, Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, Johor, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2021.02.02.049
Received 04 July 2021; Accepted 19 September 2021; Available online 30 October 2021

Abstract: This project is about the detection and localization of an object, implemented using MATLAB simulation software, in order to replace work that requires human vision and produce higher throughput and better efficiency. This project aims to develop a system that uses a built-in camera of a laptop as an input to detect an object, identify the class of the object and locate the position of the object in a real-time video. The system will go through three stages of development which are, training preparation, training process, and performance evaluation. A random number of images will be fed into the system to evaluate the performance of the system, and the result obtained will be classified into three classifications (TP, FP, and FN). More input data used to train a system will produce better accuracy and performance. The surface area of an object is another key factor that helps the system to increase its performance.

Keywords: Object Detection, Localization System, Artificial Intelligence

1. Introduction

Computer vision consists of several sub-domain and object detection, object recognition, and object localization are a few sub-domains that will be discussed in this paper. Real-time object detection can be done using image segmentation. Image segmentation uses a method, grayscale level, to divide or separate an object into several regions based on the object's characteristics [1]. A visual-based object detection, recognition, tracking, or localization is widely used in most industries due to its high speed, high precision, and most important, non-contact characteristic [2]. Another factor that a computer vision-based system is widely used is, one system is able to replace a few workers and produce better productivity and efficiency. Furthermore, human vision has limitations while a computer vision system can be trained to have a vision that surpasses a human vision. Computer vision is implemented in the field such as speech recognition, communication, autonomous vehicle, medical, security, manufacturing, etc.

Corresponding author: xavier@uthm.edu.my

The hardware used in this project to test or execute the system is a built-in laptop camera while the software used to develop and train the system in MATLAB. This project aims to develop a system that is able to detect a custom object from a real-time video. The system is also expected to locate the desired object in the real-time video frame with a bounding box and identify the class of the detected object.

Object detection and object recognition are sub-branch of computer vision and both identify an object using the same technique, however, both technique differs in terms of execution [3]. An object recognition aims at discriminating the type of object by extracting or learn discriminative features in order to classify them while object detection identifies the desired object and locate the desired object from an image frame or video sequence by surrounding the desired object with a bounding box and displays the class of the object. In other words, object detection is the combination of object recognition and object localization [4].

2. Materials and Methods

There are 3 stages to go through in order to train and develop a custom object detection model. The stages are preparation of training, the process of training, and evaluation of the trained detection model.

2.1 Training preparation

A software name MATLAB, invented by Cleve Moler, is the main platform used to produce a custom object detection model. There are several toolboxes in MATLAB that are able to aid in training and developing a detection model. The toolboxes are:

• Computer Vision Toolbox
• Deep Learning Toolbox
• Image Processing Toolbox

A folder is created directly in Local Disk (D:) in order to better organize the image gathered, training dataset, etc. Prepare a training image containing the desired object as much as possible. These images will be labeled using an app in MATLAB name Image Labeler. There are two classes of the object that had been gathered, which are pen and wallet. Other than that, a neural network architecture is also required to train a detection model. The architecture used is a modified network ResNet50.

2.2 Training process

An app name Image Labeler is used to label the region of interest (ROI) in all the images gathered. The images are separated into 2 categories, pen, and wallet, and labeled separately in a different session. The information of the ROI in each session is called a dataset. It is extracted and saved properly in a folder. The dataset will be used as training data to train a custom object detection model. In order to increase the number of a dataset without increasing the number of images in a folder, the existing dataset is augmented. This helps to improve the network accuracy. Data augmentation will randomly flip an image, randomly scale an image, jitter the color of the image or altogether.

2.3 Detection model performance evaluation

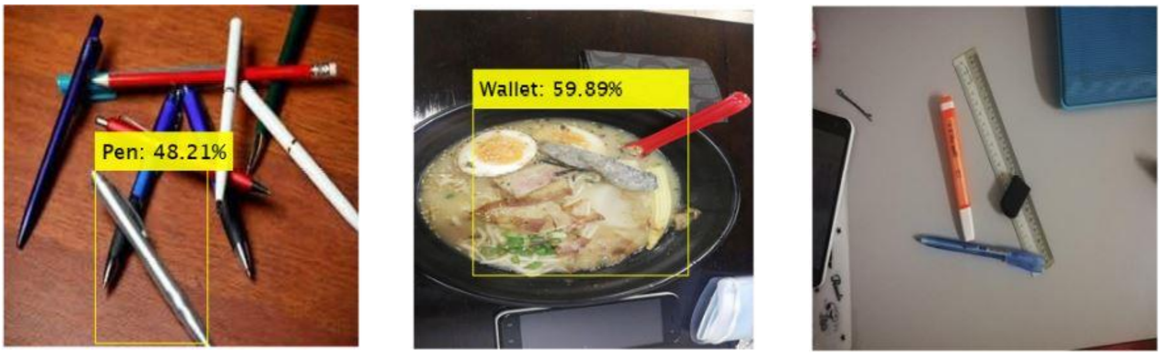
The performance of a custom object detection model is evaluated based on its precision and sensitivity. The precision and sensitivity are calculated as below:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad Eq. 1$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad Eq. 2$$

where TP is the number of true positive results, FP is the number of false positive results and FP is the number of false negative results.

According to Figure 1, it can be seen clearly the difference between all three result classifications. True positive (TP) is defined as the correct prediction of the desired object, false positive (FP) is defined as the incorrect prediction of the desired object while a false negative (FN) is defined as incapable to predict the desired object.



(a) (b) (c)

Figure 1: Result classification of (a) true positive (b) false positive and (c) false negative

3. Results and Discussion

There are a total of 208 images gathered, 60 images of a wallet, and 208 images of a pen. 18 images consisting of a wallet and 62 images consisting of a pen are randomly selected to be excluded from the training dataset and will be used to test the performance of the custom object detection model. Two custom object detection models are developed which are the wallet detection model and pen detection model

Table 1 displays the result of the wallet detection model and pen detection model with a threshold of 0.5. The result is classified accordingly to TP, FP, and FN. It is obvious that the wallet detection model has better performance than the pen detection model. This is due to the number of FN of wallet detection model is 0 while pen has 46 images classifies as FN. This greatly reduces the sensitivity of a detection model. This statement can be supported by the result in Table 2.

Table 2 displays the performance of the wallet detection model and pen detection model in terms of precision and sensitivity. The performance of the custom object detection model can be obtained by inserting the value in Table 1 into Eq. 1 and Eq. 2 accordingly. The precision and sensitivity of a detection model are calculated in ratio, hence, the closer the result is to the value of 1 the better it is. It can be seen clearly that the wallet detection model has better performance than the pen detection model.

Logically, the higher the number of images is used fed to train a custom detection model, the higher the performance. However, this logic does not apply to the custom object detection model of a pen. The factor that causes this to occur is the surface area and the position of the pen in an image. The pen is a line shape liked object, hence, if a pen is slanted in an image, the area of the ROI will increase, causing the unrelated background to be involved and this causes difficulties to the neural network to extract and learn the features of a pen correctly as the ROI is forever changing. Indirectly, this causes the performance of the pen detection model to be poor, as shown in Table 2.

Table 1: Classification result of a custom object detection model

Classification	Number of images	
	Wallet	Pen
True positive	10	8
False positive	8	8
False negative	0	46

Table 2: Performance of custom object detection model

Performance	Detection model	
	Wallet	Pen
Precision	0.5556	0.5
Sensitivity	1	0.1481

4. Conclusion

In a summary, a custom object detection model is developed successfully. The model is able to detect, identify the class of the object and locate the desired object in a given frame with a box surrounding it.

Acknowledgment

The authors would like to thank the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] Z. Loncová, L. Hargaš, D. Koniar, M. Hrianka, and A. Simonová, "Parameters measurement of the object in a video sequence," *IFAC-PapersOnLine*, vol. 28, no. 4, pp. 211–214, 2015, doi: 10.1016/j.ifacol.2015.07.034

[2] H. Yu, Y. Huang, D. Zheng, L. Bai, and J. Han, "Three-dimensional shape measurement technique for large-scale objects based on line structured light combined with industrial robot," *Optik (Stuttg).*, vol. 202, no. August 2019, p. 163656, 2020, doi: 10.1016/j.ijleo.2019.163656

[3] S. Sultania, "RealTime Object detection using YOLO and OpenCV | by Shantam Sultania | Analytics Vidhya | Medium," *Analythics Vidhya*, 2020. https://medium.com/analytics-vidhya/realtime-object-detection-using-yolo-and-opencv-8380e5537f2a

[4] C. Ambika, "Top 8 Algorithms For Object Detection," *Analytics India Magazine*, 2020. https://analyticsindiamag.com/top-8-algorithms-for-object-detection/