Penerbit Akademia Baru

Journal of Advanced Research Design

Journal homepage:
https://akademiabaru.com/submit/index.php/ard
ISSN: 2289-7984

Three-Dimensional Object Recognition using Mirrors and Deep Learning

Amiruddin Arif Saifuddin[1], Tan Lit Ken[1,*]

[1] Department of Mechanical Precision Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 29 December 2021
Received in revised form 28 April 2022
Accepted 16 May 2022
Available online 27 May 2022

Keywords:
3D; object recognition; deep learning; mirror configuration

ABSTRACT

Current object recognition model can only perceive an object from a single perspective. This may hamper its performance since some objects have a more distinct feature on one side and not the other. Therefore, past researchers rely on other methods to increase the performance from a single perspective viewpoint with additional equipment. However, this method is cost ineffective as it requires the usage of high-grade equipment and time to retrain the deep learning model. In this paper, we proposed the usage of a mirror configuration for an object recognition model to increase the chances of an object to be detected using existing trained object recognition model. It was found that the object recognition model with mirrors had an increased sensitivity of 9 %. Furthermore, the optimum mirror setup was established.

1. Introduction

Modern day society relies heavily on machine learning technology, from internet search engine to social networks websites to recommendation on e-commerce websites and is becoming increasingly present in many affordable consumer products such camera and smartphone [13]. They are used in object recognition to identify objects in images, match news article on newspaper websites, and to learn shopping behavior of customers [13,20]. Therefore, we can be certain that machine learning will play a crucial and be more heavily tied with human in the future. An example of an output using an object recognition model is shown below in Figure 1.

The advancement in deep learning in recent years has rekindle the interest in object detection and recognition. A few decades ago, it would be an impossible task for computer to differentiate and recognize any object such as a person, animal, vehicle and ordinary household item [1]. But in the last decade, many researchers were able to make significant progress in this field due to many improvements in terms of algorithm and computer hardware [13]. This is seen with many improvements in performance of numerous models in recent years [27]. For the first time, self-driving autonomous vehicle are no longer a thing of dreams but something that can be obtain.

* *Corresponding author.*
E-mail address: tlken@utm.my

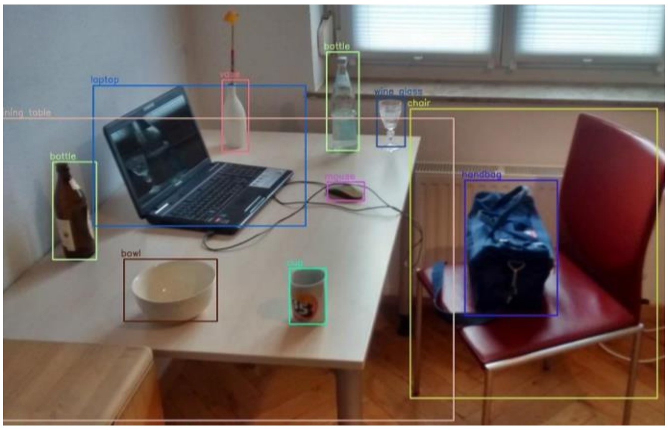


Fig. 1. Object detection example using YOLOv3

Early computer vision research can be traced back to the 1960s. One of the earliest research done was on how to match 2D features that was extracted from an image to a 3D representation of the objects [12]. This was a first step in the right direction as many researchers at that time believes that research on alphabetical character recognition will be led to a better improvement in object recognition in the future.

Object recognition software has been implanted in many industries. One of the earliest successful adaptions in the industries is in the assembly and verification process of the semiconductor industries, where complex semiconductor such as wafers is manufactured [6]. Considering the complexity of this task, human workers would have not been able to properly solve such problems reliably and efficiently. This demonstrates the economic implication that object recognition has on a country`s economy. Other fields that benefit from early implementation of object recognition system is in biomedical research and food industry [7,10].

Most of the object recognition model are only able to receive a single perspective of an object. Therefore, they do not have any information that can be acquired from other perspective. For example, certain object might have a very distinct feature at one side and not the other. Furthermore, the accuracy in current object recognition model relies heavily on the orientation of the object in an image. This shows that the current method for object recognition is not robust and flexible. In addition to that, recent advancement in the field of object recognition is heavily focused on the algorithm aspect of deep learning [27]. This can be seen with an abundance of new algorithm and backbone that is being develop in recent years. Most notable algorithm model is YoloV4 and EfficientDet [23]. The training of each model requires the training of the algorithm in order to achieve a high confidence score. In which this will require large amount of computer power [3].

Researchers have also tried other method on increasing the efficiency and accuracy by adding additional information for the algorithm, some of the researchers have tried integrating device such as Lidar or RGB-D camera to provide extra information [19]. But this has added more complexity to the object recognition model. Therefore, instead of focusing on solemnly developing newer algorithm. A mirror setup would provide a substantial alternative in creating a robust object recognition without using additional camera and sensors.

Most of the object recognition model are only able to receive a single perspective of an object. Therefore, they do not have any information that can be acquired from other perspective. For example, certain object might have a very distinct feature at one side and not the other. Furthermore, the accuracy in current object recognition model relies heavily on the orientation of the object in an image. This shows that the current method for object recognition is not robust and flexible. In addition to that, recent advancement in the field of object recognition is heavily focused on the algorithm aspect of deep learning [27]. This can be seen with an abundance of new algorithm and backbone that is being develop in recent years. Most notable algorithm model is YoloV4 and EfficientDet [23]. The training of each model requires the training of the algorithm in order to achieve a high confidence score. In which this will require large amount of computer power [3].

Researchers have also tried other method on increasing the efficiency and accuracy by adding additional information for the algorithm, some of the researchers have tried integrating device such as Lidar or RGB-D camera to provide extra information [19]. But this has added more complexity to the object recognition model. Therefore, instead of focusing on solemnly developing newer algorithm. A mirror setup would provide a substantial alternative in creating a robust object recognition without using additional camera and sensors. The objectives of the study are to develop an object recognition model using multiple mirror setup as well as evaluate the developed mirror setup performance with performance of existing object recognition model and to propose an optimum mirror setup for an object recognition model. The scope for this research is to recognize small household item or items that are 20 cm X 20 cm x 15 cm (length x width x height) in size.

This study will help improve the understanding of the object recognition model. As the need for an accurate object recognition is on high demand and increasingly required. In addition, this research may lead to other method of improving object recognition model instead on focusing on the algorithm and model training.

2. Literature Review

2.1 Object Recognition

As we are all aware, computer is not able to perceive the world around us. Therefore, a field called 'Computer Vision' was introduced to enable a computer to gain a high-level understanding from digital images of videos [2]. From an engineering standpoint, it`s aim is to replicate the human vision capability or surpass it in some task [9]. Object Recognition is a sub- domain of computer vision. It can be considered as a key output of deep learning and machine learning algorithms. Future technology such as driverless vehicle relies heavily on the ability of object recognition model to differentiate between a person, other vehicles or a stationary object such as lamp post. Other applications of object recognition in the industry such as industrial inspection and robotic vision will also increase its reliance in object recognition [6]. Figure 2 shows how current object recognition models are only able to predict a general prediction of an animal but are not able to accurately recognize the breed.

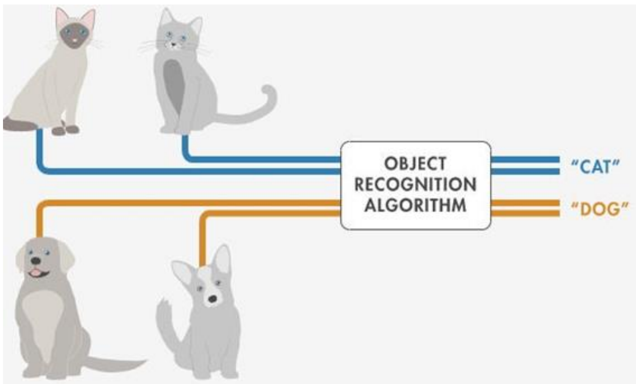


Fig. 2. Different categories of objects being identify using object recognition [21]

2.1.1 Deep learning

Deep learning techniques is one of the most used method for doing object recognition. The most used models are Convolutional Neural Network (CNN), a basic understanding on how a CNN works is that it uses the features that it has learned to identify the object. There are normally two approaches to perform object recognition using deep learning as shown in Figure 3 below.

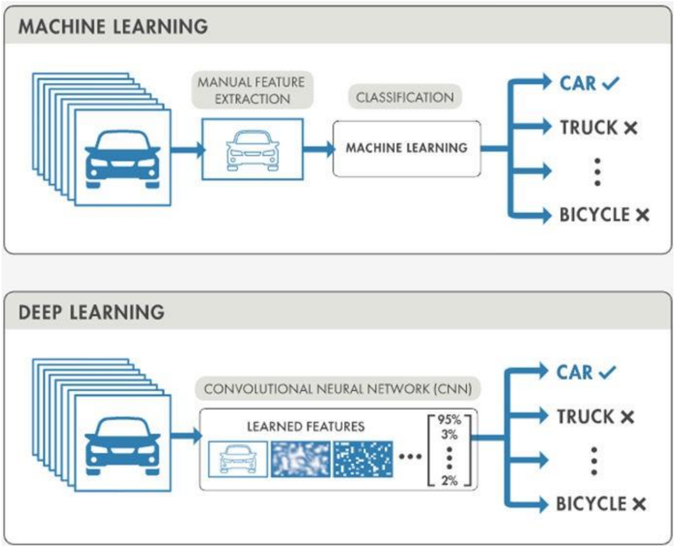


Fig. 3. Methods to develop object recognition model [3]

2.1.2 Training of a new model from scratch

To train a deep learning network from scratch requires a lot of time, substantial amount of labelled data sets and a high computational ability in order to design a network architecture that they will learn the features and build the model. This was representative when the first deep learning algorithms that produce a decent result took over a week to train on a single Graphic Processing Unit (GPU) [4]. Two ways of improving the learning rate is by splitting the workload to multiple GPU or by simplifying the feature extraction by model [4]. Hence, the training of new deep learning model is not recommended for this project as it will increase the time taken and it require a high-grade equipment.

2.1.3 Use existing pretrained deep learning model

The most used approach to deep learning is by fine-tuning a pre-trained model. This is done by taking model such as AlexNet or Darknet and training new classes by adding new data. This way, it will not only reduce the time but also improving the data for future works [22]. Furthermore, the model can directly be used if required. Therefore, the use of pre-trained deep learning is suitable for this research as it provides a quick and easy approach in the testing of object recognition performance.

2.2 Object Detection Models

There are two parts in a modern detector, a backbone and a head. A backbone is the pre-trained model that have been trained, e.g. ResNet or Darknet. The selection of the backbone can vary for different type of platform or use case since some backbones are more suitable for certain type of detection. The type of platform mention are GPU or CPU platform. The head, which is also refereed as the dense prediction is the part which is used to predict the classes and bounding boxes of objects. The head can be categorized into two kinds, which is one- stage object detector and two-stage detector as depicted in Figure 4. A short summary would be that one stage detector is more efficient and simpler, but two-stage detector have an advantage in accuracy but a more complicated pipeline in the workflow [17]. YOLOv4 is a one-stage detector, that is able to achieve good performance when compared to other standalone object recognition model [3].

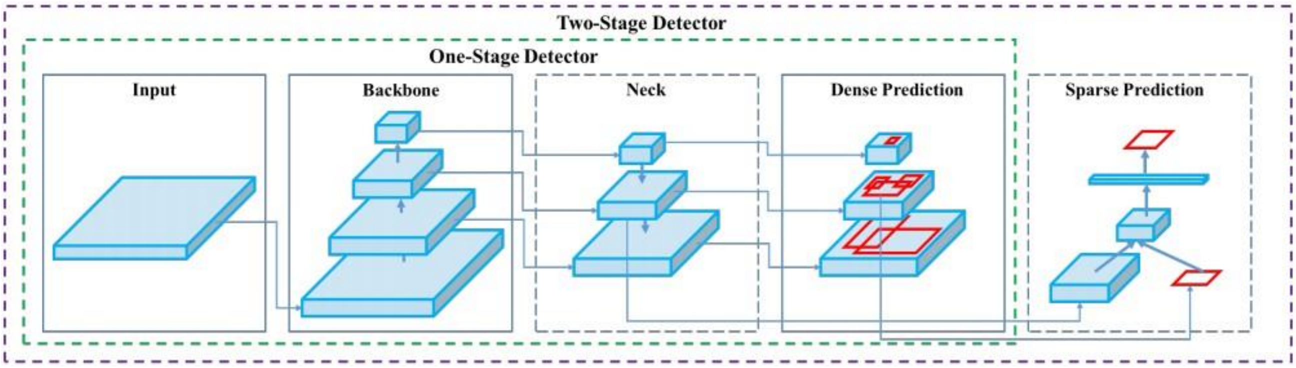


Fig. 4. Object detection workflow [3]

2.3 Datasets

Ever since the research of computer visions begins, research datasets have played an important role. They provide a means of training for new algorithms and a way for researcher to evaluate their algorithms. A large dataset is usually the preferable way to train your model as done by ImageNet [11]. But in recent years, researcher was able to utilize model that was trained using a small dataset with equal in performance to a dataset that is train using a large dataset [16]. Although, MS COCO (Figure 5) does not have as much images as other datasets such as Open Source by google, it is still able to perform great when match with the appropriate CNN.

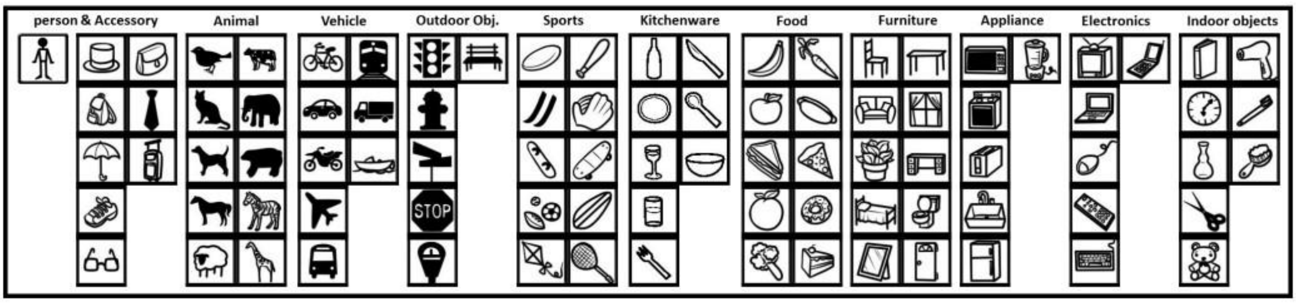


Fig. 5. MS COCO dataset that consist of 11 super-categories that are split into 91 categories [15]

2.4 Known Methods of Improving Object Recognition Model

Researchers have tried other methods to improve the performance of object recognition model from a single perspective viewpoint with additional equipment. One the most promising research is with the additional of Lidar or RGB-Depth camera into an object recognition system [19]. In addition, some research relies on extracting information such as color from an image to improve the performance of object recognition model [8]. However, these methods have still not able to provide 360 degrees view of an object and still relies from a single perspective.

2.5 Mirror in Object Recognition

Even though, mirror has always been seen as an issue in object recognition model, as they can affect the performance of a numerous vision task such as depth prediction and object detection [14]. Recently, researchers have tried to design a model that can differentiate and detect the presence of mirror inside an image [25]. This is done to separate the real and the reflection of an object which hampers certain object recognition model task, such as counting of objects. There has been work in improving computer vision using mirror, where the reflection of the mirror is used to detect an object as it pass in front of a camera with promising results [24]. Hence, the use of mirror in object recognition model is proven to have a positive effect.

2.6 Object Recognition in Industry

Object recognition system has been implanted in the industries since the early 1970s. They can do task that were almost impossible to do with the naked eye. In addition, they provide the necessary precision and accuracy that was needed for the mass production of numerous products, ranging from semiconductor, TVs and transistor [6]. This computer vision equipment was a premium towards company as they were expensive. Current object recognition system is more affordable which as help researcher to find new and exciting use case for object recognition model, e.g. the use of object recognition model in locating fruits on trees [10]. Hence, the use of object recognition model in other fields such as recycling, or quality control should be looked.

3. Methodology

The aim of this research is to develop an object recognition model using multiple mirror setup. Firstly, the image of objects is taken using the mirror setup. Then, the image is insert into YOLOv4 to get the confidence score of objects. From the average result of confidence score, the accuracy of mirror setup is determined. Lastly, the accuracy of the mirror and non-mirror setup is then compared to identify which model gives the better accuracy.

3.1 Data Collection

To begin with this research, a mirror setup had to be built to ease the changing of mirror inclination. A GoPro camera (model Hero8 Black) was used as the camera for collecting the images. The distance of camera was fixed at 60 cm from the reference point O (Figure 6). While the placement of object is 20 cm from reference point. The type of object will be diversified to get an accurate representation. A total of 342 images were taken, 4 images are taken for each object at different setup. The image taken will form two group sets of data. First, the data that will be used to test mirror configuration, while the other set are used for non-mirror configurations. The ambient lighting was kept the same throughout the collection of the images. Figure 7 shows the top view of setup and Figure 8 to 10 depicts the cell phone C with different mirror inclination and angles between the mirror.

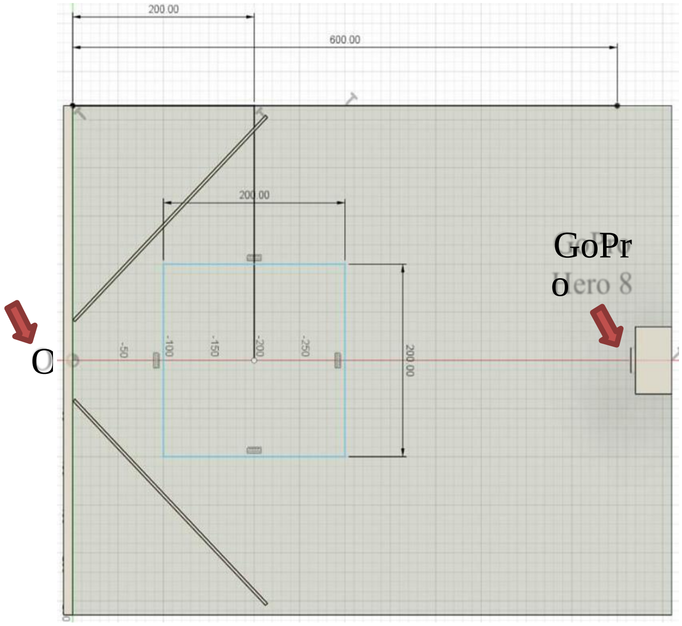


Fig. 6. Layout dimension and placement of equipment

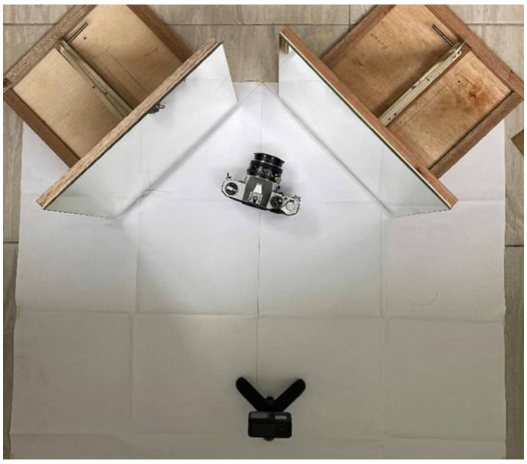


Fig. 7. Top view of setup

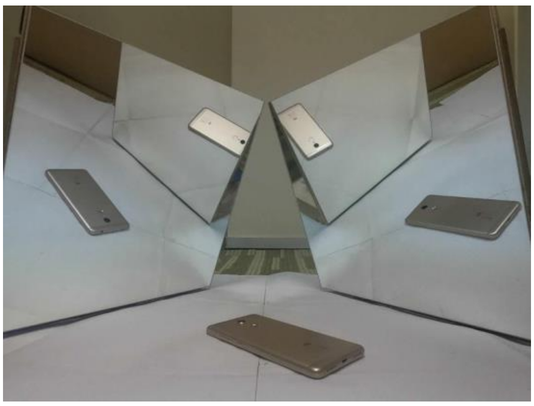


Fig. 8. Cell phone C (Mirror inclination = 70˚, Angle between mirror = 70˚)

Fig. 9. Cell phone C (Mirror inclination = 80˚, Angle between mirror = 70˚)

Fig. 10. Cell phone C (Mirror inclination = 90˚, Angle between mirror = 100˚)

3.2 Confidence Score

By utilizing YOLOv4, we obtained a confidence score for the images that was taken (Figure 11 to 13). The number of successful detections was observed for every image. Some images showed multiple confidence score results while some does not give any result. Performance of the configurations and setup was calculated using the highest confidence score in each image. The data was then recorded in Microsoft Excel to be analyzed. The example shown below are the result of the image that were taken. In cases where they are extra reflection on the mirror as shown in Figure 14 only the confidence score on the two main reflections on the mirror was taken.

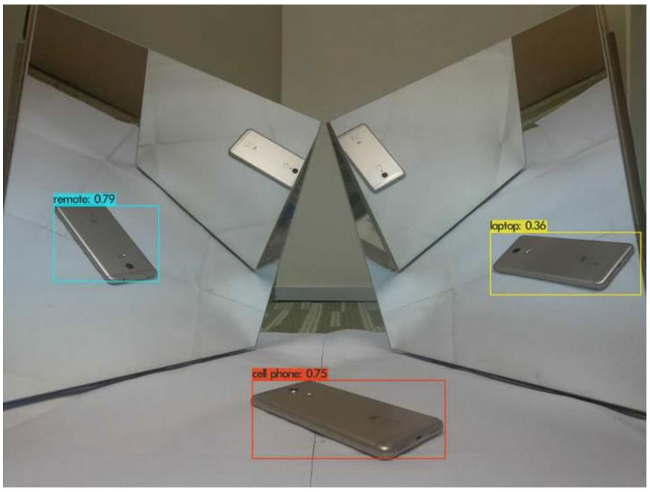


Fig. 11. Cell phone C using confidence score using YOLOv4 (Mirror inclination = 70˚, Angle between mirror = 70˚)

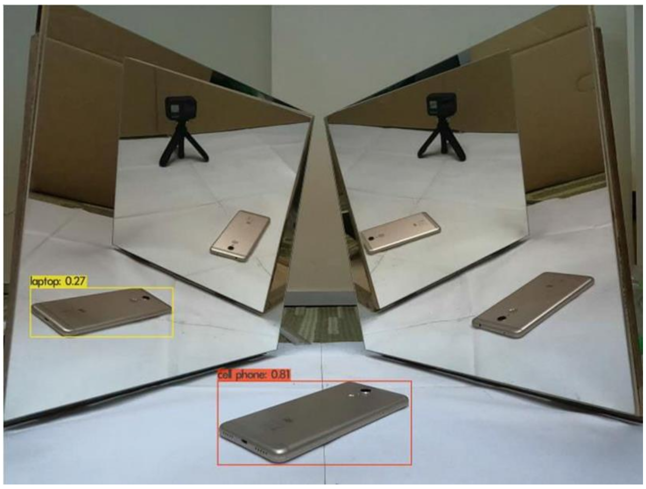


Fig. 12. Cell phone C using confidence score using YOLOv4 (Mirror inclination = 80˚, Angle between mirror = 70˚)

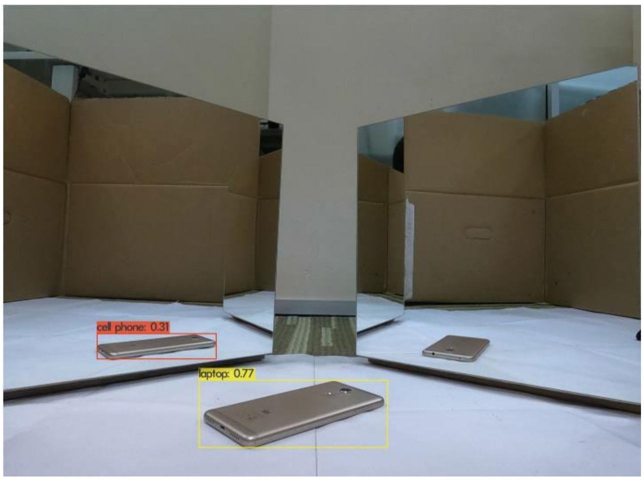


Fig. 13. Cell phone C using confidence score using YOLOv4 (Mirror inclination = 90˚, Angle between mirror = 100˚)

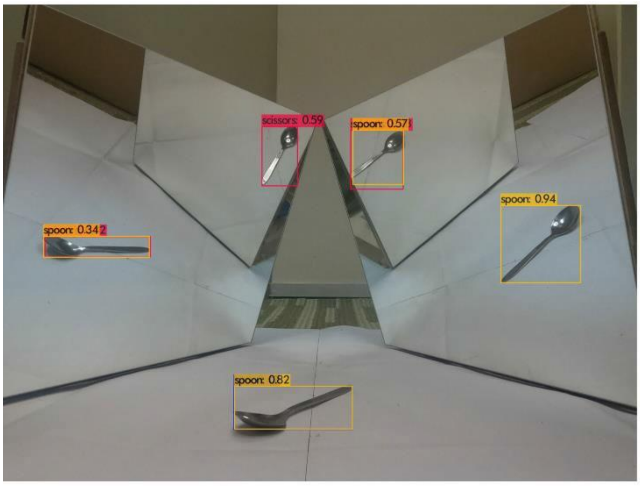


Fig. 14. Multiple extra result on mirror

3.3 Receiver Operating Characteristics (ROC) Curve

Only the highest confidence score in each image was recorded. This was done as it is intuitive for a system or a person to decide based on the highest confidence value. This is to prioritize and evaluate the setup performance solemnly on its ability to detect objects. The collected data is then input into an algorithm that will create a ROC curve (Figure 15) to determine a suitable threshold for the performance comparison. A ROC curve is a graph showing the performance of a classification model at different classification threshold [26].

Classification threshold in object recognition is known as decision threshold. Decision threshold is crucial in identifying at which threshold does the object recognition is at its optimal performance [18]. The two main variables of ROC curve is the True Positive Rate and False Positive Rate.

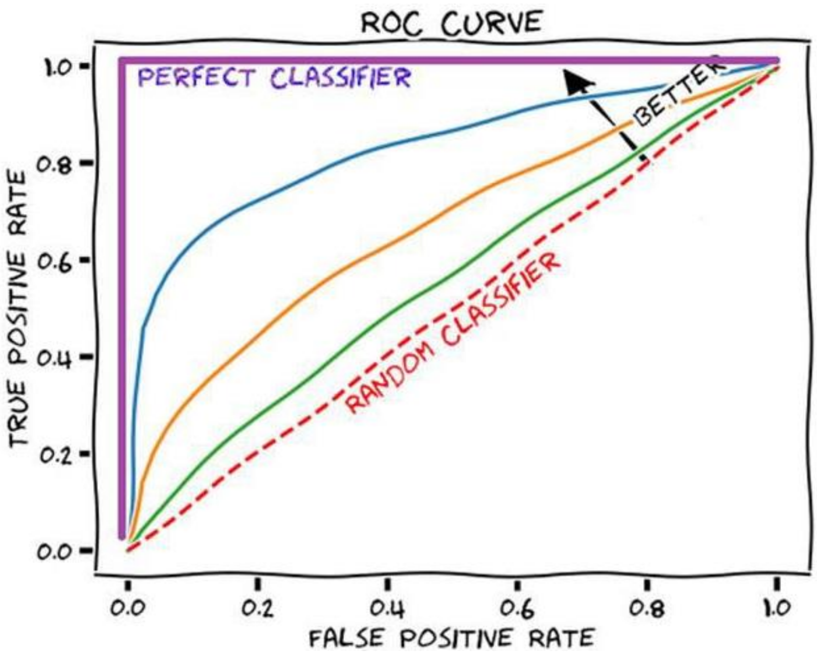


Fig. 15. ROC curves [5]

3.4 Confusion Matrix

After collecting the data and finding the decision threshold using ROC curve depicted in Figure 15, the data was then classified into 4 main classifications. The classifications were:

i. True Positive
ii. False Positive
iii. False Negative
iv. True Negative

The classification was derived from 4 distinctive scenarios. These scenarios can only be classified after the value of decision threshold is established as in Table 1. Table 2 then list the confusion matrix used for the classification.

With the identifications of each outcome done, the performance of the object recognition with the mirror can be found. Figure 16 provides a visual representation of the confusion matrix.

Table 1
Possible scenario based on the decision threshold

Decision threshold: 0.60	
Scenario 1 Actual : Cup Confidence Score : 0.80 Prediction : Cup	Scenario 3 Actual : Cup Confidence Score : 0.70 Prediction : Book
Scenario 2 Actual : Cup Confidence Score : 0.40 Prediction : Cup	Scenario 4 Actual : Cup Confidence Score : 0.35 Prediction : Book

Table 2
Confusion matrix used for classification

		Prediction	
		True	False
Accept	Yes (Confidence Score > Decision Threshold)	True Positive • Prediction: Cup (True) • Accept: Yes • Scenario 1	False Positive • Prediction: Book (False) • Accept: Yes • Scenario 3
	No (Confidence Score < Decision Threshold)	False Negative • Prediction: Cup (True) • Accept: No • Scenario 2	True Negative • Prediction: Book (False) • Accept: No • Scenario 4

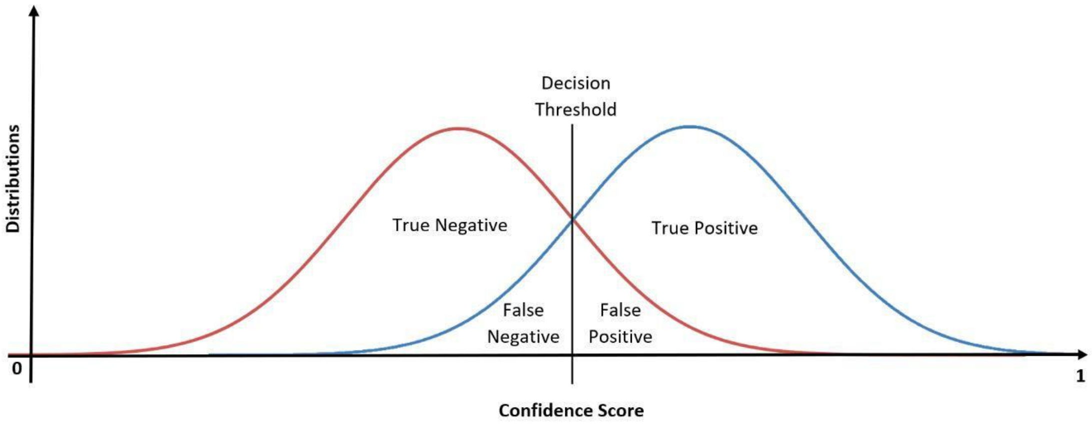


Fig. 16. Visualization of the 4 main conditions

3.5 Performance Calculation

Upon completing the confusion matrix, the performance of the mirror configurations can be determined through identifying the configuration that have the highest precision and recall value. Precision can be identified as the proportion of positive detections that is actually correct while recall is the proportion of actual positives that was identified correctly.

The definition of precision (Figure 17) and recall (Figure 18) can be better understood by visualizing the proportion in the figures below,

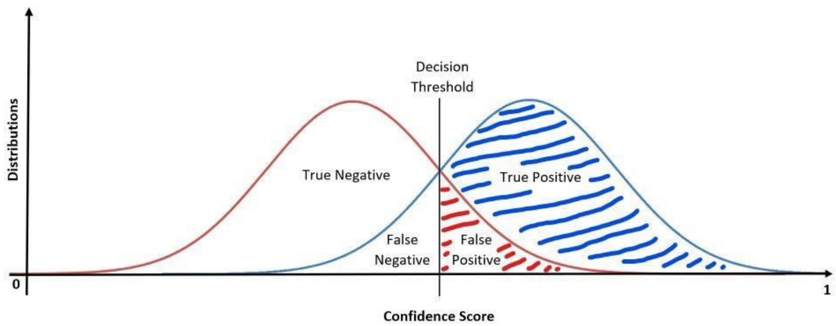


Fig. 17. Visualization of precision

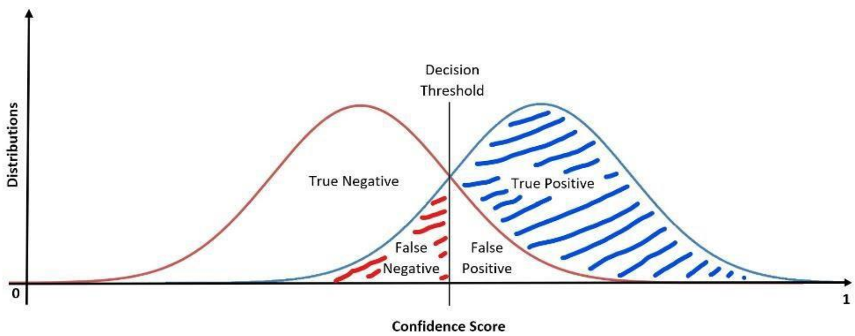


Fig. 18. Visualization of recall

Accuracy can also be used to calculate the performance of an object recognition performance, but it is not preferable if the data is heavily imbalanced. For example, where a significant disparity between the number of positives and negative labels occurs. Furthermore, a well-trained object recognition model should not be able to detect an object wrongly. Since it shows that the system is flawed, as a result a decision threshold is required to evaluate performance of object recognition model.

Since most of the detection that are wrongly predict occurs with a low confidence score. Therefore, it can be classified as a True Negative in order to fulfil the condition of confusion matrix.

4. **Results and Discussion**

Firstly, the confidence score of each detection using mirror and non-mirror was obtained. Then the decision threshold was achieved using ROC curve. Subsequently, the confusion matrix was established using the decision threshold that was set using the ROC curve. Then, the value of precision and recall was calculated and the comparison between the mirror and non-mirror configurations was made.

Figure 19 and 20 show the distribution of the confidence score for the mirror setup and non-mirror setup respectively. The object recognition with mirror setup has a higher frequency of successfully detecting object when compared to the non-mirror setup, with 171 correct detections. However, the non-mirror setup was only able to detect 145 correct detections.

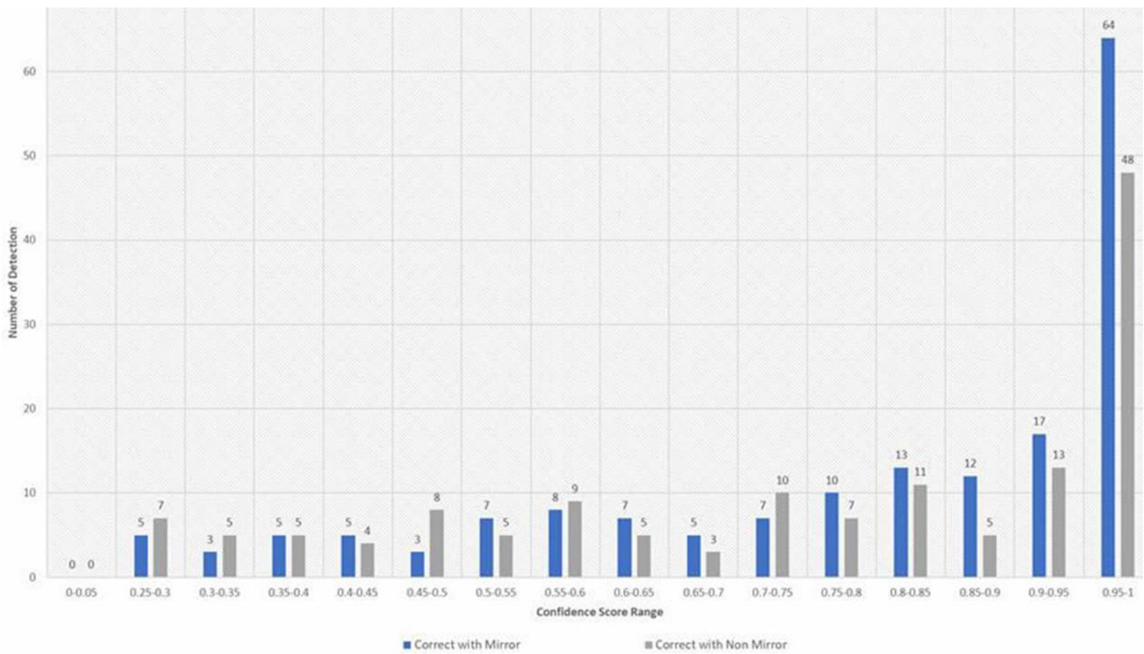


Fig. 19. The number of correct detections using mirror and non-mirror

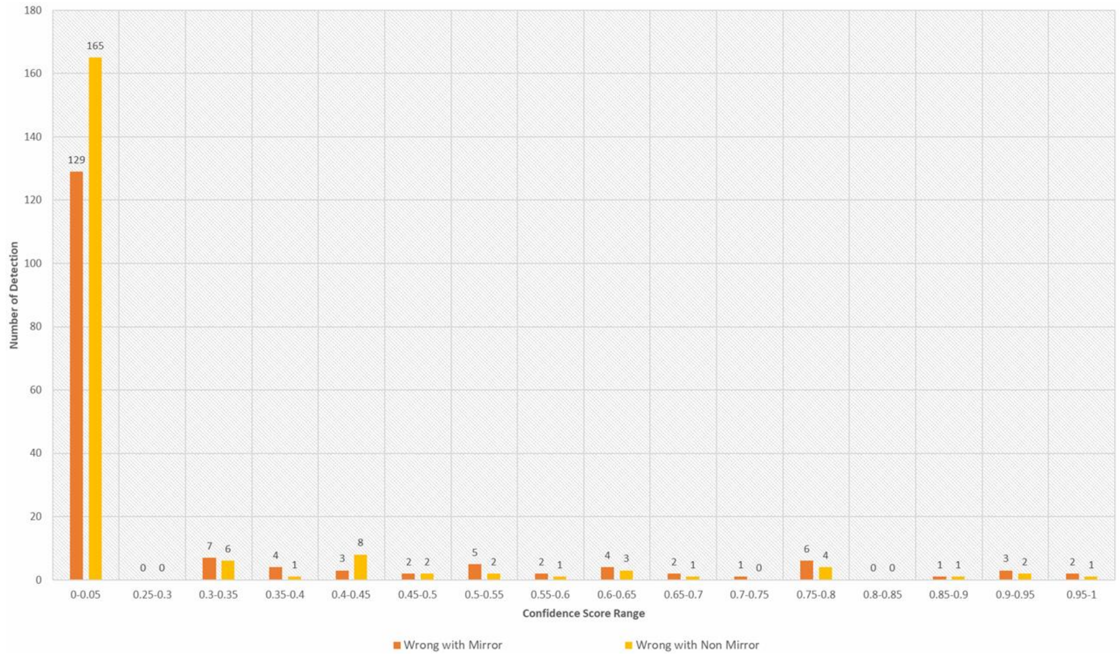


Fig. 20. The number of wrong detections using mirror and non-mirror

4.1 ROC Curves

As mentioned in the methodology, the ROC curve was used in order to determine the decision threshold for the object recognition model. Figure 21 was the fitted ROC curve for both the mirror and non-mirror configurations.

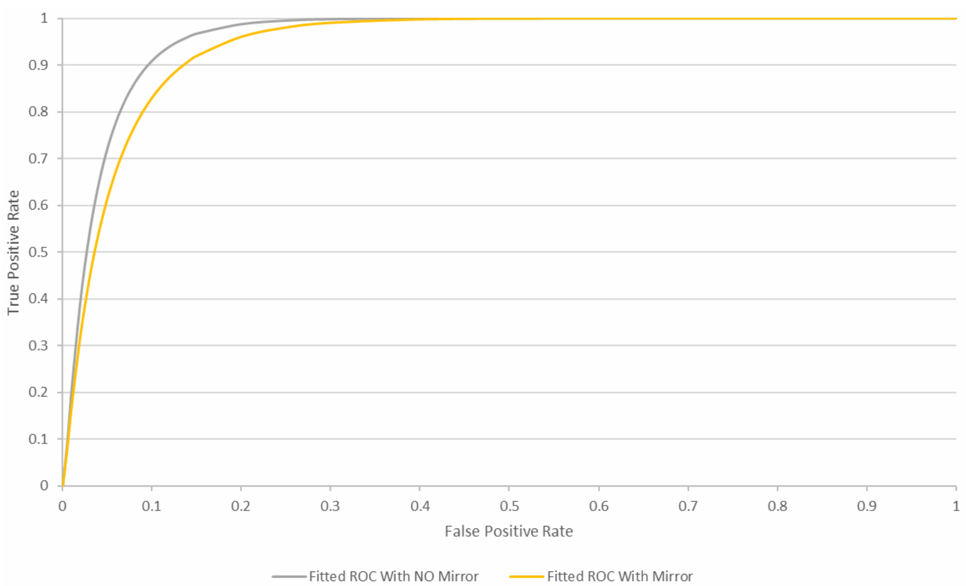


Fig. 21. Fitted ROC curve for both configurations

4.1 Performance Comparison

This section discussed the comparison that was done to evaluate the performance of the object recognition model. Firstly, a comparison between a mirror and non-mirror was done to determine the performance that was achieved with the additional of mirror. Next, a comparison between all of the different setup was done to determine which setup was the most suitable for an object recognition model.

4.1.1 Comparison between mirror and non-mirror configurations

Based on Table 3, the recall value of mirror setup was significantly better than the non-mirror setup with a 9 % improvement while maintaining the similar precision performance. This improvement in recall corresponds with the increase in the number of correct detections as highlighted in Figure 19 and a lower number of wrong detections as shown in Figure 20. The similarity in the precision was a result of using the same image for both the mirror and non-mirror configurations. In other words, the result is expected since they are from the same image that was taken.

Table 3
Precision and recall for mirror and non-mirror setup

Configuration	Precision	Recall (Sensitivity)
Non-Mirror	0.89	0.70
Mirror	0.88	0.79

4.1.2 Comparison between different mirror setups

Table 4 shows the calculated performance of different setup in term of their precision and recall. From Figure 22, it can be concluded that only one setup was able to achieve a recall value of more than 0.9, which is the setup with an inclination of 70 degrees and the angle between mirror of 70 degrees. The same figure also provided the worst setup for an object recognition model with a setup of an inclination of 90 degrees and an angle between mirror of 100 degrees.

Table 4
Precision and recall for different mirror setup

Mirror Inclination (˚)	*70*				*80*				*90*			
Angle Between Mirror (˚)	*70*	*80*	*90*	*100*	*70*	*80*	*90*	*100*	*70*	*80*	*90*	*100*
Precision	0.87	1	0.85	0.78	0.92	0.87	1	0.84	1	1	0.67	0.5
Recall	0.93	0.73	0.85	0.58	0.75	0.81	0.84	0.84	0.88	0.71	0.67	0.67
Accuracy	0.91	0.88	0.88	0.77	0.81	0.82	0.91	0.81	0.92	0.83	0.75	0.71

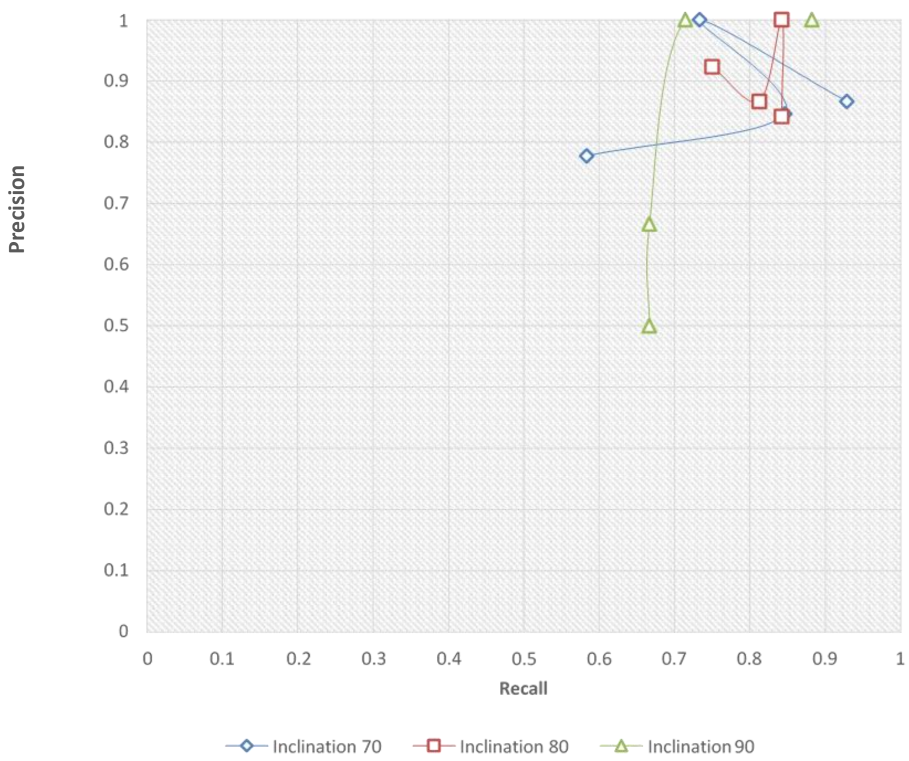
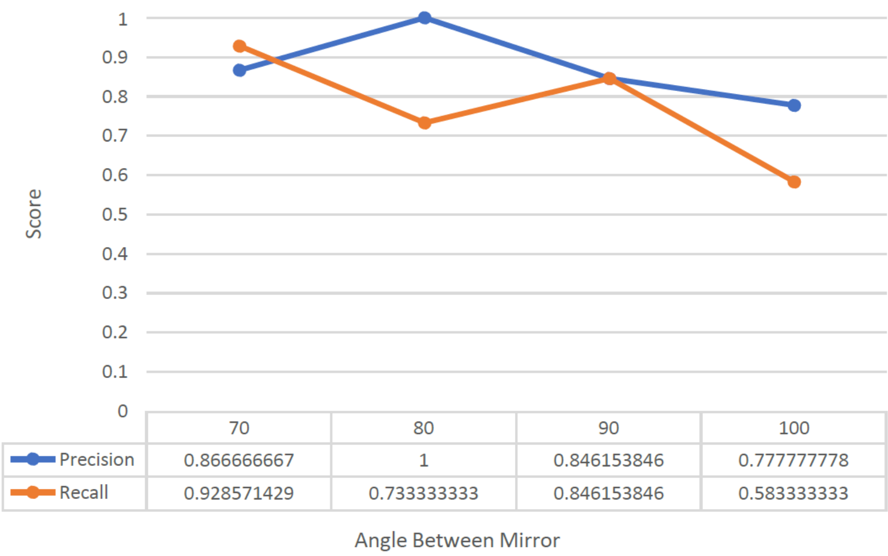


Fig. 22. Precision against recall

The trend in Figure 23 to 25 below shows that as the angle between mirror increased from 70 degrees to 100 degrees, the precision and recall value for both of the figure shows a downward trend. While in both the precision and recall value does not show any significant changes with the increase in angle between mirrors.



	70	80	90	100
Precision	0.866666667	1	0.846153846	0.777777778
Recall	0.928571429	0.733333333	0.846153846	0.583333333

Fig. 23. Performance of setup with mirror inclination 70 with different angle between mirrors

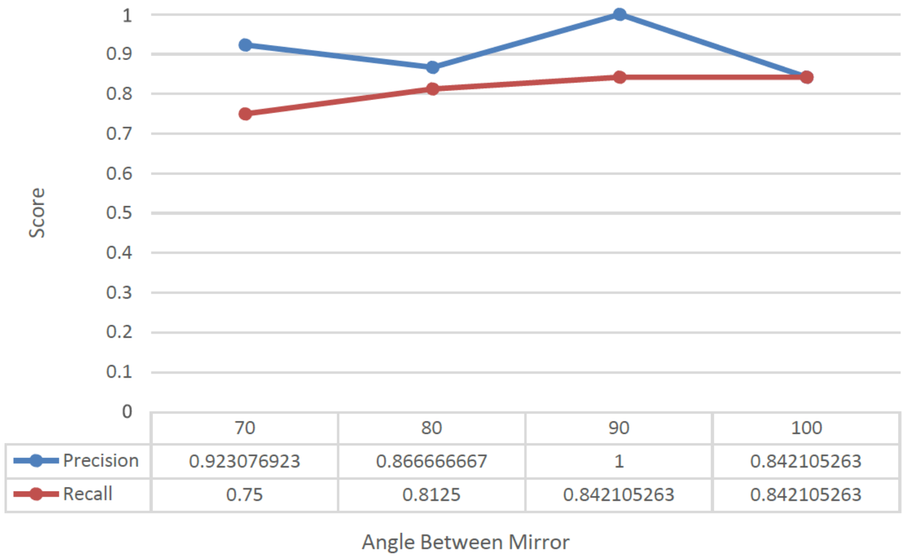


Fig. 24. Performance of setup with mirror inclination 80 with different angle between mirrors

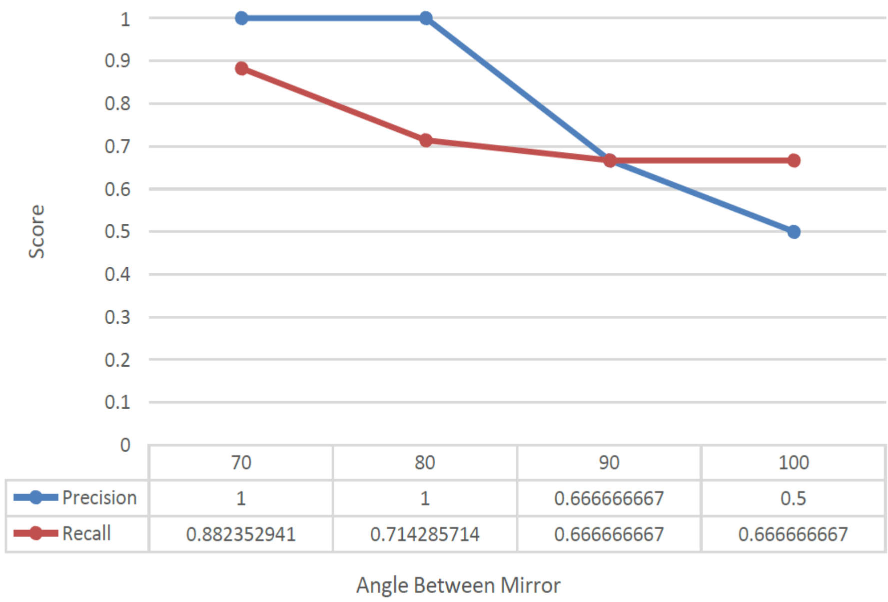


Fig. 25. Performance of setup with mirror inclination 90 with different angle between mirrors

5. Conclusions

To improve the robustness of object recognition model, we proposed the use a mirror setup on object recognition system. The experiment on 342 images with mirror setup yield promising result. With an increase of 9 % in sensitivity by using a mirror, we were able to validate the claim that using a mirror is a viable method to increase the performance of object recognition model without the need for expensive equipment.

It is shown that with the additional of mirror, the object recognition model can perform better than a stand-alone object recognition model and certain mirror setup can produce better performance than others.

A comparison between suggested mirror setup has been made to determine the best setup and the best setup is found to be a system with a mirror inclination of 70 degrees and an angle between mirror of 70 degrees. This setup was able to achieve the highest recall (sensitivity) out of all the setup that was suggested while also having a respectable precision value.

References

[1] Andreopoulos, Alexander, and John K. Tsotsos. "50 years of object recognition: Directions forward." Computer vision and image understanding 117, no. 8 (2013): 827-891. https://doi.org/10.1016/j.cviu.2013.04.005
[2] Ballard, Dana Harry, and Christopher M. Brown. Computer vision. Prentice Hall Professional Technical Reference, 1982.
[3] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934 (2020).
[4] Chahal, Karanbir Singh, Manraj Singh Grover, Kuntal Dey, and Rajiv Ratn Shah. "A hitchhiker's guide on distributed training of deep neural networks." *Journal of Parallel and Distributed Computing* 137 (2020): 65-76. https://doi.org/10.1016/j.jpdc.2019.10.004
[5] Draelos, Rachel. "Measuring performance: Auc (auroc)." *Glass Box* (2019).
[6] Ejiri, Masakazu. "Machine vision in early days: Japan's pioneering contributions." In *Computer Vision–ACCV 2007: 8th Asian Conference on Computer Vision, Tokyo, Japan, November 18-22, 2007, Proceedings, Part I 8*, pp. 35-53. Springer Berlin Heidelberg, 2007. https://doi.org/10.1007/978-3-540-76386-4_3
[7] GALLUS, GIUSEPPE, and GIUSEPPE REGOLIOSI. "A decisional model of recognition applied to the chromosome boundaries." *Journal of Histochemistry & Cytochemistry* 22, no. 7 (1974): 546-553. https://doi.org/10.1177/22.7.546
[8] Gevers, Theo, and Arnold WM Smeulders. "Color based object recognition." In *Image Analysis and Processing: 9th International Conference, ICIAP'97 Florence, Italy, September 17–19, 1997 Proceedings, Volume I 9*, pp. 319-326. Springer Berlin Heidelberg, 1997. https://doi.org/10.1007/3-540-63507-6_217
[9] Huang, Thomas S. "Computer vision: Evolution and promise." *CERN European Organization for Nuclear Research-Reports-CERN* (1996): 21-26.
[10] Jimenez, A. R., R. Ceres, and Jose L. Pons. "A survey of computer vision methods for locating fruit on trees." *Transactions of the ASAE* 43, no. 6 (2000): 1911-1920. https://doi.org/10.13031/2013.3096
[11] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." *Communications of the ACM* 60, no. 6 (2017): 84-90. https://doi.org/10.1145/3065386
[12] Langdon, J. "The perception of three-dimensional solids." *Quarterly Journal of Experimental Psychology* 7, no. 3 (1955): 133-146. https://doi.org/10.1080/17470215508416686
[13] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521, no. 7553 (2015): 436-444. https://doi.org/10.1038/nature14539
[14] Lin, Jiaying, Guodong Wang, and Rynson WH Lau. "Progressive mirror detection." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3697-3705. 2020. https://doi.org/10.1109/CVPR42600.2020.00375
[15] Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. "Microsoft coco: Common objects in context." In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pp. 740-755. Springer International Publishing, 2014. https://doi.org/10.1007/978-3-319-10602-1_48
[16] Liu, Shuying, and Weihong Deng. "Very deep convolutional neural network based image classification using small training sample size." In *2015 3rd IAPR Asian conference on pattern recognition (ACPR)*, pp. 730-734. IEEE, 2015. https://doi.org/10.1109/ACPR.2015.7486599
[17] Lu, Xin, Quanquan Li, Buyu Li, and Junjie Yan. "Mimicdet: Bridging the gap between one-stage and two-stage object detection." In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16*, pp. 541-557. Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-58568-6_32
[18] Mandrekar, Jayawant N. "Receiver operating characteristic curve in diagnostic test assessment." *Journal of Thoracic Oncology* 5, no. 9 (2010): 1315-1316. https://doi.org/10.1097/JTO.0b013e3181ec173d
[19] Maturana, Daniel, and Sebastian Scherer. "Voxnet: A 3d convolutional neural network for real-time object recognition." In *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 922-928. IEEE, 2015. https://doi.org/10.1109/IROS.2015.7353481
[20] Najafabadi, Maryam M., Flavio Villanustre, Taghi M. Khoshgoftaar, Naeem Seliya, Randall Wald, and Edin Muharemagic. "Deep learning applications and challenges in big data analytics." *Journal of big data* 2 (2015): 1-21. https://doi.org/10.1186/s40537-014-0007-7
[21] Object Recogntion: 3 things you need to know. (n.d.).

[22] Redmon, J. Darknet: Open Source Neural Networks in C (2016).

[23] Tan, Mingxing, Ruoming Pang, and Quoc V. Le. "Efficientdet: Scalable and efficient object detection." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10781-10790. 2020. https://doi.org/10.1109/CVPR42600.2020.01079

[24] Wu, Jing, and Ze Ji. "Seeing the unseen: Locating objects from reflections." In Towards Autonomous Robotic Systems: 19th Annual Conference, TAROS 2018, Bristol, UK July 25-27, 2018, Proceedings 19, pp. 221-233. Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-96728-8_19

[25] Yang, Xin, Haiyang Mei, Ke Xu, Xiaopeng Wei, Baocai Yin, and Rynson WH Lau. "Where is my mirror?." In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8809-8818. 2019. https://doi.org/10.1109/ICCV.2019.00890

[26] Zou, Kelly H., A. James O'Malley, and Laura Mauri. "Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models." *Circulation* 115, no. 5 (2007): 654-657. https://doi.org/10.1161/CIRCULATIONAHA.105.594929

[27] Zou, Zhengxia, Keyan Chen, Zhenwei Shi, Yuhong Guo, and Jieping Ye. "Object detection in 20 years: A survey." Proceedings of the IEEE 111, no. 3 (2023): 257-276. https://doi.org/10.1109/JPROC.2023.3238524