sensors

Article

Optimization of Identification Trash on The House Compound Using Convolutional Neural Network (CNN) and Sensor System

Emil Naf'an [1,2], Riza Sulaiman [1] and Nazlena Mohamad Ali [1,*]

[1] Institute of IR4.0, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia; P88912@siswa.ukm.edu.my
[2] Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, Padang 25221, Indonesia
* Correspondence: …@ukm.edu.my

Abstract: This study aims to optimize the object identification process, especially identifying trash in the house compound. Most object identification methods cannot distinguish whether the object is a real image (3D) or a photographic image on paper (2D). This will be a problem if the detected object is moved from one place to another. If the object is 2D, the robot gripper only clamps empty objects. In this study, a Convolutional Neural Network (CNN) combination with a LiDAR (Light Detection and Ranging) sensor was carried out with an accuracy of ±2mm. After testing 11 types of trash on four CNN architectures (AlexNet, VGG16, GoogleNet, and ResNet18), the accuracy results are 93.4%, 92.6%, 97.5%, and 95.9%. This result is perfect for object identification. However, it needs to be optimized using a LiDAR sensor to determine the object in 3D or 2D. It will be ignored if the fast scanning process with the LiDAR sensor detects non-real (2D) trash. If Real (3D), the trash will be scanned in detail to determine the gripper's position in lifting the trash. The time efficiency generated by fast scanning is between 13.33% to 59.26% depending on the object's size. The larger the object, the greater the time efficiency obtained. In conclusion, optimization using a combination of CNN and LiDAR Sensors can identify trash objects correctly and determine whether the object is real (3D) or not (2D), so a decision may be made to move the trash from the detection location.

Keywords: Optimization; Identification; Trash; Convolutional Neural Network (CNN); Sensor

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Sensors* **2022**, *22*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

1. Introduction

A clean house compound is a dream for everyone. Trash can be caused by leaves falling from trees or plastic waste such as plastic bottles, snack packets, and others. Usually, the trash on the house compound is cleaned by the owner or the person assigned. If homeowners are busy and do not have time to clean the house compound, then the house will look unclean because the trash is scattered in the house compound.

Currently, the state considers essential urban services, such as water, sanitation and solid waste management, to be the responsibility of local or national governments [1]. Research in the identification and classification of waste has been done. However, it was not optimal regarding the amount of trash detected and accurate trash detection. At the same time, optimization plays an essential role in computer vision because many computer vision algorithms employ an optimization step at some point in their proceeding [2]. One of them was done by Y.Fuchikawa in 2005 about trash collection using an OSR robot (Outdoor Service Robot) [3]. Unfortunately, this research is only to collect plastic waste in PET bottles. Another research was conducted by an Italian research group led by Barbara Mazzolai [4]. They named the robot DustCart. This robot groups junk types based on user input. After the user inputs the type of trash, the robot opens the trash store according to the type of trash input. However, robot cameras are for navigation, avoiding obstacles, and not classifying trash types.

Advancing improvement in reliability and processing speed of the vision system, and experiment with other trash must be done [3]. Improvement was conducted at Stanford University in which research the type of waste has been classified using the CNN (Convolution Neural Network) method [5]. However, the amount of trash that can be identified only up to 6 items of trash alone, while in the house compound, the trash more than 6 items.

Identification of the waste is an important step before separation and it can be done efficiently with the help of different machine learning and image processing algorithms. Convolutional neural network (CNN) are most preferred for classification of images [6]. However, most methods of identifying objects cannot distinguish whether the object is a real (3D) image or an image in the form of a photograph on paper (2D). In Figure 1 there is a test object (mouse) in 2D and 3D. The object is detected using Convolutional Neural Network (CNN). The test results show that 2D and 3D (real) objects can be detected well with a prediction of 98%. In Figure 1 the 2D object is located on the right side of Figure 1. While the 3D object is located on the left side of Figure 1.

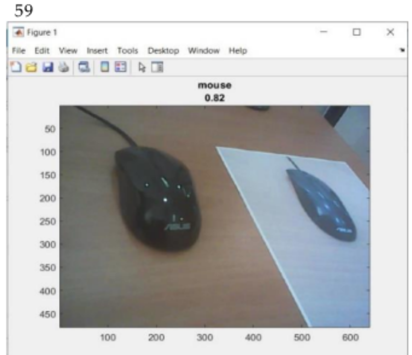
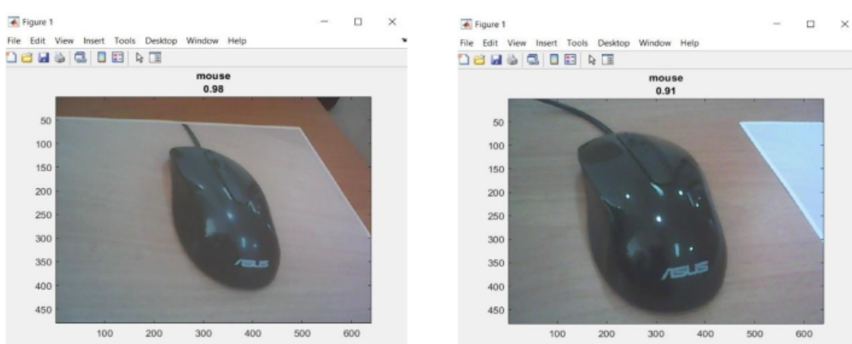


Figure 1. A sample of detection object (mouse) in 2D and 3D.

Then, each object is detected, namely 2D objects and 3D objects using the Convolutional Neural Network (CNN). Hardware using an Intel Core i7 (8th gen) laptop, 16GByte RAM, NVIDIA GeForce RTX2060 graphics card with MATLAB 2020a software. For 2D objects, samples are used in the form of photos of objects on paper. The test results for 2D objects can be seen in Figure 2 (a). The test results are obtained with a prediction of 98%.



(a) (b)

Figure 2. A sample of detection object (mouse). (a) Photo of mouse on paper (2D); (b) Real mouse photo (3D).

Furthermore, testing for 3D objects (real). This object is placed next to the 2D object. The test results can be seen in Figure 2 (b). The test results are obtained with a prediction of 91%. From these three test results, it can be said that the Convolutional Neural Network (CNN) can detect objects well. This will be a problem if the object is moved from one place to another using gripper of the robot. How would it be if the robot wants to move the trash (object), while the trash (object) was just a photograph of trash on paper? Therefore, it requires a robot that can identify/recognize trash (objects) in real (3D). In any application, we can use a particular classifier, and try to optimize its performance. The usual approach is to try several different classifiers and choose the one that performs the best on a separate validation set (G. Dougherty, 2013). In this case, we propose Convolutional Neural Network (CNN) combined with sensors to identify whether the object is real or not. CNN was chosen because at present time CNN is the best machine learning method in object identification. Therefore, to optimize identification of trash in house compound using Convolutional Neural Network (CNN) and Sensor System is proposed.

2. Materials and Methods

In this study, 2D and 3D image data were taken using an IP Camera. The image data is processed by a laptop using MATLAB R2020a software. Data collection is done by using the trash data itself. This data is grouped into 11 types of data, namely; Cardboard, Fabric, Food Packaging, Fruit, Glass, Leaf, Metal, Paper, Plastic, Rubber, and Wood.

(a) (b)

Figure 4. (a) Trash Data (Own Dataset). (b) 11 types of trash data.

The data were tested with 4 Convolutional Neural Network (CNN) architectures, namely: AlexNet, VGG16, GoogleNet, and ResNet18. Block diagram of the proposed system can be seen in Figure 5.

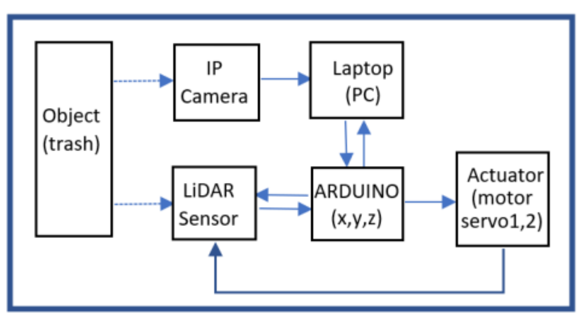


Figure 5. Block Diagram

After the image data is captured by the IP camera, MATLAB program resizes the image. This is useful for further processing on CNN Image Detection. Basically the CNN architecture used is capable of classifying up to 1000 types of objects. However, to speed up the process of classifying trash objects, modifications are made to the feature learner. The original value is 1000 fully connected layers changed to 11 fully connected layers. This value is changed on every architecture, both AlexNet, VGG16, GoogleNet, and ResNet18.

Furthermore, the MATLAB program ensures that the image is a trash object. If the object is not trash, it is ignored. If it is trash, the laptop (PC) sends a command to Arduino to activate the actuator (servo1 and servo2). This command is based on the detected object bounding box value. The actuator moves the LiDAR sensor in X and Z coordinates based on the value. After that, the data from LiDAR sensor is sent by Arduino to the Laptop (PC). The proposed flowchart can be seen in Figure 6.

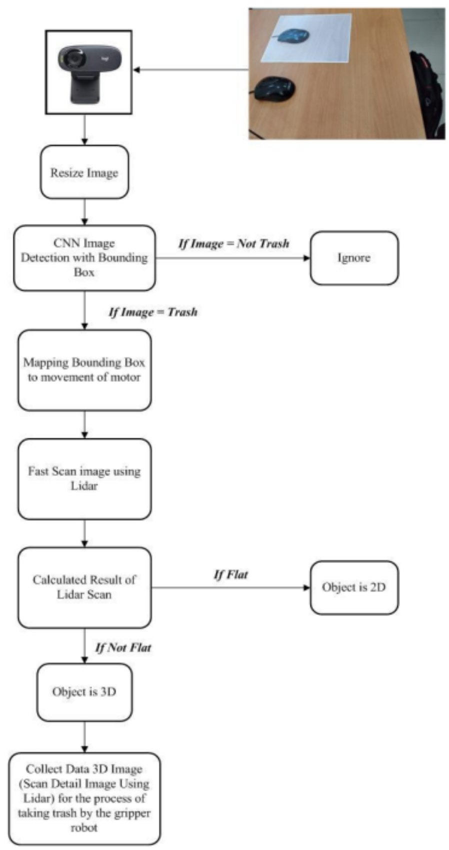


Figure 6. Flowchart of the proposed study.

The Mapping Bounding box is used to convert the movement of the servo motor to match the size of the object to be detected. After getting the value of x and z, it is calculated/mapped to the degree of motor servo movement. In the scanning process, servo motor 1 (x) and servo motor 2 (z) move at the boundary in the bounding box.

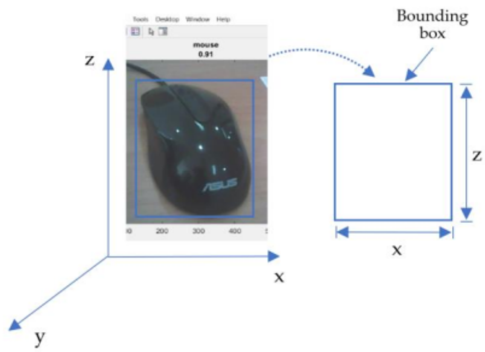


Figure 7. Convert bounding box to x,y,z. coordinates.

In the designed system, the LiDAR sensor is mounted on top of a servo motor 1. The scanning process is then carried out using a LiDAR sensor. The scanning process is carried out in 2 stages, namely the fast scanning process and the detailed scanning process. This is for the efficiency of scanning time. If the detected object is flat, then the object is a 2D object. If it's 2D then the detailed scanning process does not need to be done. However, if the object is not flat, then the detected object is a 3D object. If the object is 3D then the detail process needs to be done. It aims to determine with certainty the geometric shape of the object. With the object's geometry data, the robot gripper is easier in the process of lifting trash objects.

2.1.TF40 LiDAR

TF40 is an mm-level accuracy LiDAR with a range up to 40m. TF40 has the following features: Features: High accuracy, Tiny, Small light spot, Visible laser, easier for aiming. Table 1 is the main parameter of TF40 and Figure 8 shows the physical form of the TF40 LiDAR sensor and its dimensions.

Table 1. Main Parameters of TF40.

Parameter Name		**Standard Version**
Product performance	Range	0.04-40m 90% reflectivity, 0.04-20m 10% reflectivity
	Accuracy	± 2 mm
	Distance resolution	1mm
	Frame rate	5 Hz
Optical parameters	Light source	LD
	Wavelength	635nm
	Laser class	CLASS 2 (EN 60825)
	Detection angle	< 1mrad
Electrical parameters	Supply voltage	3.3 V
	Average current	≤ 180mA
	Power consumption	≤0.6W
	Communication voltage level	LVTTL (3.3V)

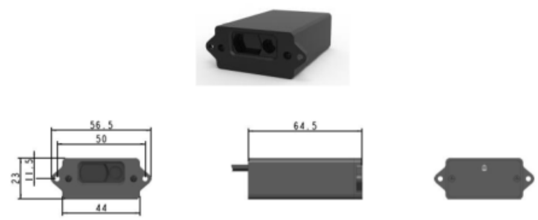


Figure 8. TF40 LiDAR.

2.2. Fast Scanning Image Using LiDAR

The scanning process is fast starting from the top left position to the bottom right. The fast-scanning process is carried out 5 times according to the path in Figure 9.

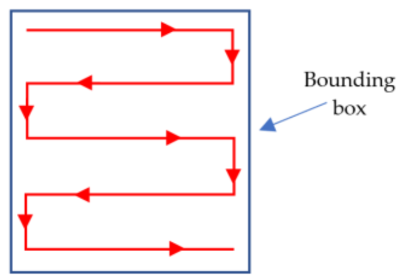


Figure 9. Fast scanning process path.

The height of the bounding box (z) can be different depending on the size of the detected trash object. However, the value of the degree of movement of the servo motor 2 (z) in this fast-scanning system is obtained from the height of the bounding box divided by 5 as shown in the following formula:

$$\text{Servo 2 (z)degree} = \frac{\text{Height of bounding box}}{5} \tag{1}$$

After getting the value of x and z based on bounding box, it is calculated (mapped) to the degree of motor servo movement. Motor servo 1 (x) and servo 2 (z) are moving like the figure 10.

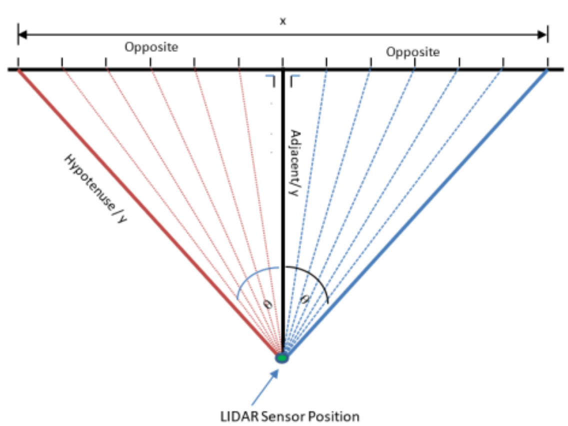


Figure 10. The position of the LiDAR sensor during the scanning process.

In Figure 10 it can be seen the position and movement of the LiDAR sensor during the scanning process. To find out that the detected object has flat, concave or convex sides, the following algorithm is used.

If sensor value = hypotenuse then
the line / point of the image is a flat plane
If sensor value > hypotenuse then
the lines / dots of the image become concave
If sensor value < hypotenuse then
the lines / dots of the image become convex

In the fast scanning process, the data read is time data (Time Stamp) and distance data (Distance). Figure 10 is an example of a graph of the LiDAR TF40 sensor reading.

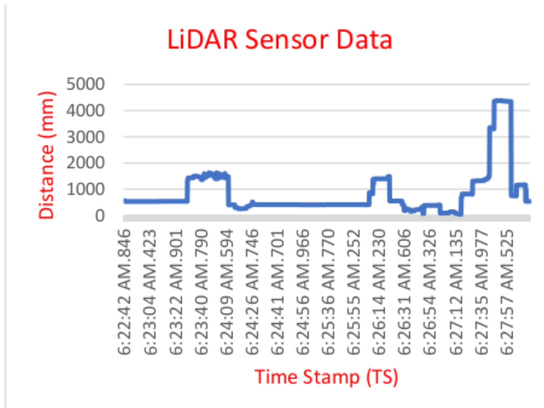


Figure 11. Example of LiDAR sensor data.

2.3. Detail Scanning Image Using LiDAR

Detail Scanning is useful for the process of taking objects using a robot gripper. If the gripper is not positioned properly, the lifting process may fail. The scanning process is almost the same as fast scanning, but the degree of movement of the servo motor has been determined from the start, which is 2 degrees. The details can be seen in Figure 12.

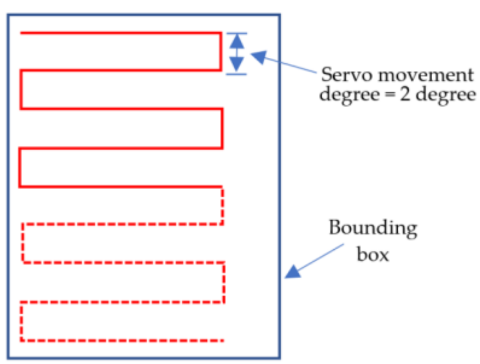


Figure 12. Detail scanning process path.

3. Results

Image data were tested for 11 types of waste using 4 CNN architectures, namely AlexNet, VGG16, GoogleNet and ResNet18. The number of images for each type of trash is 150 pieces. The data is augmented into 1200 images, so the total image is 13,200 images. The data is divided into 70% training data and 30% test data. Figure 13 is a training data test using the GoogleNet architecture.

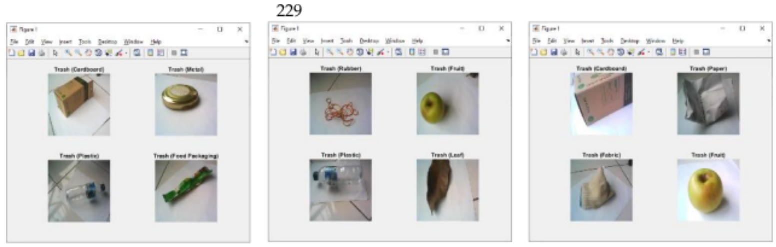
Figure 13. Result of Training progress Trash Object Using GoogleNet.

In Figure 13 the validation accuracy generated by Google Net is 93.18%. The results of the AlexNet, VGG16, and ResNet18 training progress can be seen in table 2.

Table 2. Training Progress Result.

CNN Architecture	Validation Accuracy (%)
AlexNet	84.09
VGG16	79.55
GoogleNet	93.18
ResNet18	88.64

Table 2 shows that the CNN GoogleNet architecture has the highest Validation Accuracy value, while the CNN VGG16 architectures have the lowest Validation value. After the training process is complete, the identification test of the trash object is carried out. Tests of the 11 types of waste can be seen in Figures 14. This test uses the AlexNet architecture.



(a) (b) (c)

Figure 14. Results of the trash object identification test. (a) Cardboard, Metal, Plastic, Food Packaging; (b) Rubber, Fruit, Plastic, Leaf ; (c) Cardboard, Paper, Fabric, Fruit.

3.1. Confusion Matrix for trash classification testing

This trash object identification system is tested for its performance using a confusion matrix. The four CNN architectures (AlexNet, VGG16, GoogleNet, ResNet18) were tested according to their respective architectures. Figure 15 shows the results of the confusion matrix with Google Net.

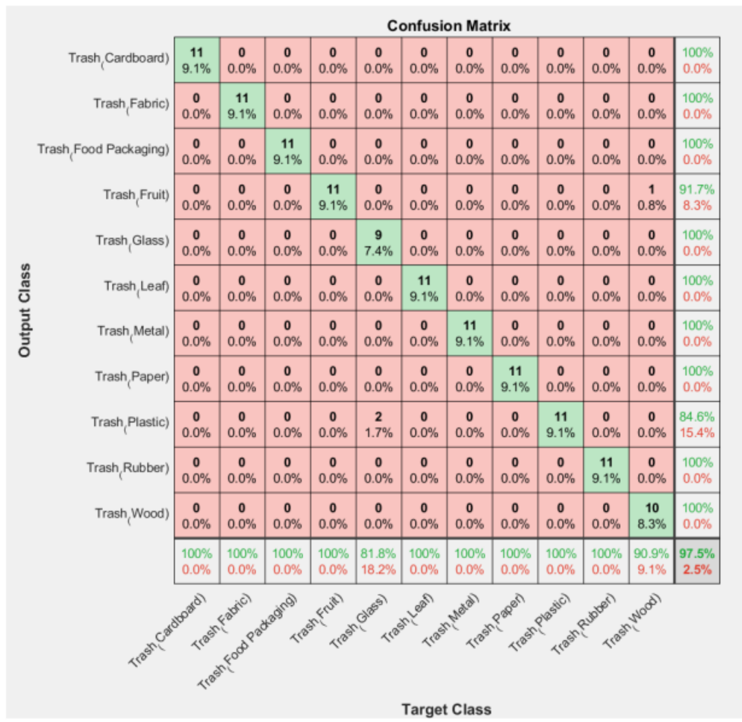


Figure 15. Result of Confusion Matrix Trash Using GoogleNet.

Table 3 is the result of comparing the accuracy of the confusion matrix. In the table it can be seen that the VGG16 architecture has the lowest accuracy value, namely 92.6%, while the GoogleNet architecture has the highest value, namely 97.5%.

Table 3. Comparison of confusion matrix.

CNN Architecture	Accuracy
AlexNet	93.4
VGG16	92.6
GoogleNet	97.5
ResNet18	95.9

3.2. Result of Fast Scanning Image

Fast scanning image is used to quickly ensure that the identified image is an image in 2D or 3D. The results of the fast-scanning image can be seen in Figure 15.

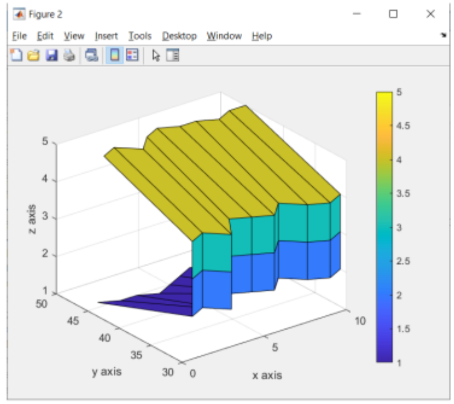


Figure 15. Result of Fast Scanning Image of the trash object (box)

In Figure 15 it can be seen that the results of the fast-scanning image divide the value of the z axis into 5. This value comes from the servo motor 2.

3.3. Result of Detail Scanning Image

After the fast-scanning process states that an observed object is an object in 3D, then the detailed image scanning process is carried out. The function of this detail scanning image is to ensure the position of the gripper when lifting trash.

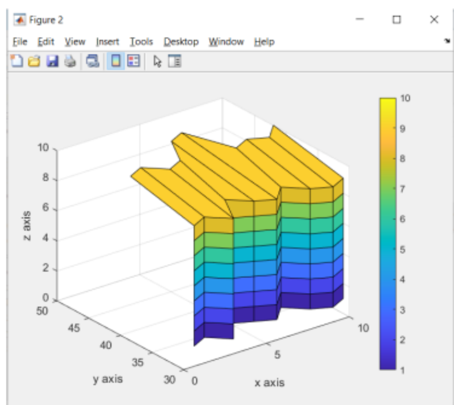


Figure 16. Result of Detail Scanning Image of the trash object (box)

In Figure 16 it can be seen that the detailed image scanning results are better than the fast-scanning image results. This is because the amount of data on the y-axis and z-axis is more than the amount of data on the y- and z-axis in fast scanning images.

3.4. Result of Time Speed Comparison between Fast Scanning Image and Detail Scanning Image

One of the optimizations carried out in the trash identification process is to make time efficiency in the process of identifying trash objects. Therefore, it is necessary to examine the difference between fast scanning images and detailed scanning images as depicted in Figure 17. From the results of this test, the efficiency of time consumption used in the trash identification process will be obtained.

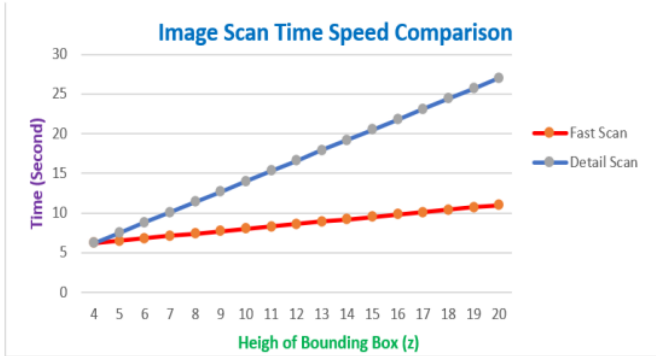


Figure 17. Result of Fast Scanning Image and Detail Scanning Image of the trash object

4. Discussion

In the initial experiment, the total images entered were 690 images with details: cardboard 50, fabric 15, food packaging 50, fruit 50, glass 50, leaf 165, metal 65, paper 100, plastic 120, rubber 15 and wood 15. In the results of training progress, accuracy ranges from 65.3% to 74.5%. Accuracy results on the Confusion Matrix ranged from 84.2% to 92.1%.

To overcome the low value of Accuracy, the number of each data image in each class needs to be added especially on images less than 20. Images of each class are added so that a minimum of 150 images per class. Then each image is augmented 7 times, so that each class has 1200 images. There is an increase in accuracy in training progress with accuracy ranged from 79.55% to 93.18%. The increase in accuracy in the Confusion Matrix also occurred ranging from 92.6% to 97.5%. Thus, the addition of data images must be done to increase the accuracy of the identification of the trash object.

Figure 10 illustrates the position of the LiDAR sensor during the scanning process. Based on the figure, there are 3 formulas used, namely :

1. If the object is straight in front of the LiDAR sensor then the y-coordinate :

$$y = \ distance\ \ value\ \ that\ \ measured\ \ by\ \ LiDAR\ \ sensor \tag{2}$$

2. If the object is on the front left side / the front right side of the LiDAR sensor, the y-coordinate value can be calculated by the formula :

$$y = \ Cos\ \ \theta\ \ \ x\ \ \ distance\ \ measured\ \ by\ \ LiDAR\ \ sensor \tag{3}$$

3. The x-coordinate value can be calculated by the formula :

$$\mathrm{x = Degree\ \ of\ \ servo\ \ Motor1 - 90} \tag{4}$$

The value of 90 is due to servo motor1 being set at 90 degrees. After all the x values are read and stored in the matrix variable, then the x value is added to the maximum x value with the formula :

$$\mathrm{x = x + maximum\ \ value\ \ of\ \ x} \tag{5}$$

This aims to make the x-coordinate values all positive.

After the x-coordinate and y-coordinate values are obtained, then the z-value is obtained by the formula :

$$z = \text{Degree of servo Motor2} - 90 \quad (6)$$

After all the z values are read and stored in the matrix variable, then the z value is added to the maximum z value with the formula :

$$z = z + \text{maximum value of } z \quad (5)$$

5. Conclusions

The optimization of the trash object detection system has been successfully carried out by using a fast-scanning system based on bounding boxes. The time efficiency obtained ranges from 13.33 – 59.26% depending on the size of the detected object. The larger the object, the greater the time efficiency obtained. bigger the object, the more time it takes. Testing is limited to objects with a size of 15 cm x 20 cm.

Testing the identification of trash objects using several CNN architectures, namely: AlexNet, VGG16, GoogleNet and ResNet18. All these architectures have a trash object idenfication accuracy of 93.4%, 92.6%, 97.5% and 95.9%, respectively. This system uses a LiDAR sensor to ensure that the object is real or not. The results of LiDAR scanning in graphic form can be produced properly, because the LiDAR sensor has a reading accuracy of 2 mm. The graph is the basis for determining the position of the gripper in lifting the trash object. Future research will focus on the trash collection process based on the resulting graph.

Author Contributions: Supervision: R.S, N.M.A.; draft preparation: E.N., R.S, N.M.A.; review and editing: E.N., R.S, N.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This paper supported the grant : GPK-4IR-2020-021

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, [E.N], upon reasonable request.

Acknowledgements: The authors would also like to thank the Editor and the anonymous reviewers for their valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] I. Baud, *Solid Waste Management and Recycling*. KLUWER ACADEMIC PUBLISHERS, 2004.

[2] M. A. Treiber, *Optimization for Computer Vision*. 2013.

[3] Y. Fuchikawa *et al.*, "DEVELOPMENT OF A VISION SYSTEM FOR AN OUTDOOR SERVICE," pp. 100–105, 2005.

[4] E. Spadoni and C. Laschi, "Focus on Social and Legal Challenges," no. March, pp. 59–67, 2011.

[5] M. Yang and G. Thung, "Classification of Trash for Recyclability Status," *CS229Project Rep.*, pp. 1–6, 2016.

[6] K. S. Hulyalkar S., Deshpande R., Makode K., "Implementation of Smartbin Using Convolutional Neural Networks," *Int. Res. J. Eng. Technol.*, vol. 5, no. 4, pp. 3352–3358, 2018.

1. Sensor citation..
2. UKM…prof riza 5 citations

Optimization

ORIGINALITY REPORT

19%	18%	10%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	repository.upiyptk.ac.id Internet Source	6%
2	Submitted to Kumoh National Institute of Technology Graduate School Student Paper	3%
3	mdpi-res.com Internet Source	2%
4	Submitted to Chosun University Student Paper	2%
5	www.irjet.net Internet Source	1%
6	archive.org Internet Source	1%
7	jtsiskom.undip.ac.id Internet Source	1%
8	elib.dlr.de Internet Source	<1%
9	journals.ums.ac.id Internet Source	<1%

10	Daniel Meissner. "Simulation and calibration of infrastructure based laser scanner networks at intersections", 2010 IEEE Intelligent Vehicles Symposium, 06/2010 Publication	<1%
11	assets.researchsquare.com Internet Source	<1%
12	Takeshi Nishida, Yuji Takemura, Yasuhiro Fuchikawa, Shuichi Kurogi et al. "Development of Outdoor Service Robots", 2006 SICE-ICASE International Joint Conference, 2006 Publication	<1%
13	Zhang, Harry, Matthew Smith, and Gerald Witt. "Driving Task Demand‚ÄìBased Distraction Mitigation", Driver Distraction Theory Effects and Mitigation, 2008. Publication	<1%
14	ebin.pub Internet Source	<1%
15	jurnal.unidha.ac.id Internet Source	<1%
16	www.pinterest.com Internet Source	<1%
17	Submitted to Roanoke College Student Paper	<1%

18	Yasuhiro Watanabe. "Development of a Sensor System for Outdoor Service Robot", 2006 SICE-ICASE International Joint Conference, 10/2006 Publication	<1 %
19	Paula M. L. F. Alvarenga, Ana P. Gonçalves, Rosa M. C. S. C. Fernandes, Amarilis P. A. de Varennes et al. "Effect of Organic Residues and Liming Materials on Metal Extraction from a Mining-Contaminated Soil", Bioremediation Journal, 2008 Publication	<1 %
20	www.internationalstudentsymposium.com Internet Source	<1 %
21	"Intelligent and Fuzzy Techniques: Smart and Innovative Solutions", Springer Science and Business Media LLC, 2021 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		