Research Article
Journal of Smart Systems Research (JOINSSR) 3(2), 120-127, 2022
Received: 3-November-2022 Accepted: 12-December-2022

SAKARYA UNIVERSITY OF APPLIED SCIENCES

YOLO – Based Waste Detection

Kenan Erin[1], Bünyamin Bingöl[2], Barış Boru[3]

[1] Mechatronics Engineering Department, Sakarya University of Applied Sciences,
[2,] Electrical and Electronics Engineering Department, Sakarya University of Applied Sciences,
[3] Mechatronics Engineering Department, Sakarya University of Applied Sciences,

ABSTRACT

The management of recycling wastes is one of the most important issues because of the increasing production rates. The collecting and recycling of waste are also becoming more crucial for economic and environmental reasons because landfill space is becoming more and more limited. Automatic sorting systems are defined as systems that separate recyclable waste materials with robotic manipulators where human intervention is minimal. In this study, while determining the type of waste, the location of the waste will be determined in 3D with a depth camera and image processing techniques.

Keywords: YOLO, RealSense, Waste Recognition, 3D Position Measurement

1 Introduction

One of the biggest issues facing developed nations is the separation of recyclable materials. Waste, particularly plastic waste, is growing daily. Increased plastic manufacture leads to some issues, including ecological, water, and air pollution. It is possible to recycle waste and lessen adverse environmental effects by practicing effective recycling.

The two categories of recycling sorting procedures are manual sorting and automatic sorting. Large recycling operations cannot use manual sorting processes, which are frequently described as workers visually identifying and classifying recyclables. Automatic separation systems are systems where human intervention is minimal and the efficiency obtained from recycling materials is high.

Systems for separating plastic, paper, metal, etc. are an essential step in the trash recycling process. A system with high accuracy will recycle better. While a badly constructed system results in significant cost losses, a well-recycled system ensures the reuse of waste and energy generation.

With the research studies carried out in recent years, reversible separation systems have an important place in the literature today. The techniques applied in waste separation and the results obtained are shown in the literature review. Studies on the existing literature can be summarized as follows. Tatzer et al. [1] carried out the detection of waste materials by the hyperspectral imaging method in their study. In the experimental setup, paper and cardboard wastes from the separation process were used. After the study, they separated paper and cardboard with a success rate of over 90%. Scavino et al. [2] developed a prototype mechanism for the automatic detection and separation of plastic bottle waste on the conveyor. They can overcome the problems of detecting low light and deformed bottles thanks to the hardware they utilized in the study and the software they developed. In the study, which was tested on 50 plastic materials, a successful classification of 97% was obtained. Ozkan et al. [3] provided image processing-based separation of recycled plastic materials in their study. Images that were taken via camera were first used to identify plastic wastes, and then various methods of image processing were used to remove noise from the image. Support Vector Machine (SVM) algorithm was used to classify plastic waste, and the successful classification rate was determined to be approximately 90%. Meng and Chu [4] carried out waste detection on the dataset in their study. The authors obtained 10108

* Corresponding Author's email: kenanerin@subu.edu.tr

images with dataset augmentation methods for 2527 images. Then, they used the SVM+HOG, Simple CNN, Resnet50, and HOG+CNN algorithms to separate this dataset into five different categories: cardboard, glass, metal, plastic, and general waste. After comparing the performance of the algorithms they had used, they came to the conclusion that ResNet50 had achieved the highest level of success (92.25%). Li and Wei [5] detected waste on the data set they collected from the camera in their study. They extracted the data set they collected to 20000 pictures with various methods. Then, with the Inception-V3 model, they classified the pictures into 4 categories as recyclable waste, household waste, hazardous waste and general waste. At the end of the study, they determined training success as 99.3% and test success as 93.2%. In their study, Jingyi et al. [6] performed waste recognition and separation with a mobile manipulator. First, they detected the garbage with the algorithm they created on the TACO data set in the literature, which they called GarbageNet. Then, in order for the mobile manipulator to separate the waste, they determined the position of the waste and the grip position on the 3D image. At the end of the study, they revealed that the waste collection process in closed areas works successfully with the mobile manipulator.

In our study, the images captured by the camera were used to detect waste in real time and identify its location. First, training was conducted for 4 different wastes, metal, paper, glass, and plastic, over the data set. Then YOLOv4 algorithm was used for real-time work in waste detection. Using the 3D camera, the location of the image in 3D and the center of the image, whose type was identified, were both determined.

2 Material and Method

In this section, the algorithm and hardwares used to solve the problem are explained.

2.1 YOLO (You Look Only Once)

The YOLOv4 algorithm was used in problem-solving. The most important reason why the YOLOv4 algorithm is preferred is that its performance is quite high compared to its others versions, as shown in Figure 1.

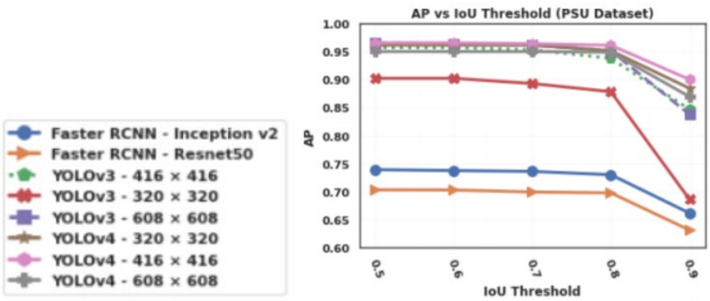

Figure 1. Comparison of YOLO Models [1]

YOLO, an object detection technique used for real-time image processing, has lately gained popularity. Figure 3 shows its general structure. The YOLO Algorithm can predict faster than other algorithms. YOLO applies a CNN (Convolutional Neural Network) to the picture, divides the picture into grids, calculates bounding boxes and the appropriate confidence score for each grid, and calculates the bounding boxes with the estimated confidence score [7]. Figure 2 shows how the rules of the Yolo algorithm limit the detection of objects in an image. These boundaries are indicated by rectangles between detected objects [8].

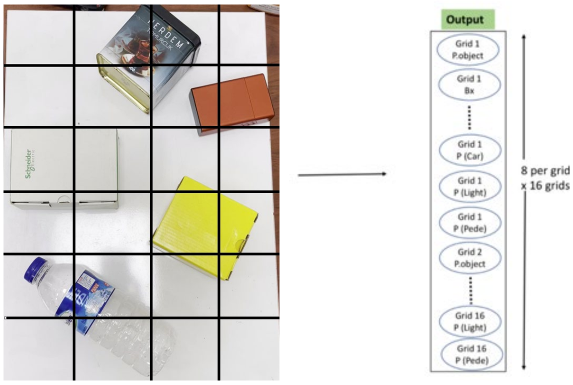

Figure 1: Structure of the YOLO algorithm

The Yolo algorithm can process images at around 40-90 FPS (Frames per second). Therefore, it is very fast compared to other methods. This shows that a video can be processed by the Yolo Algorithm in real-time with a delay of a few milliseconds [7].

Compared to another object detection method, Yolo is said to be 1000 times faster than R-CNN and 100 times faster than Faster R-CNN.

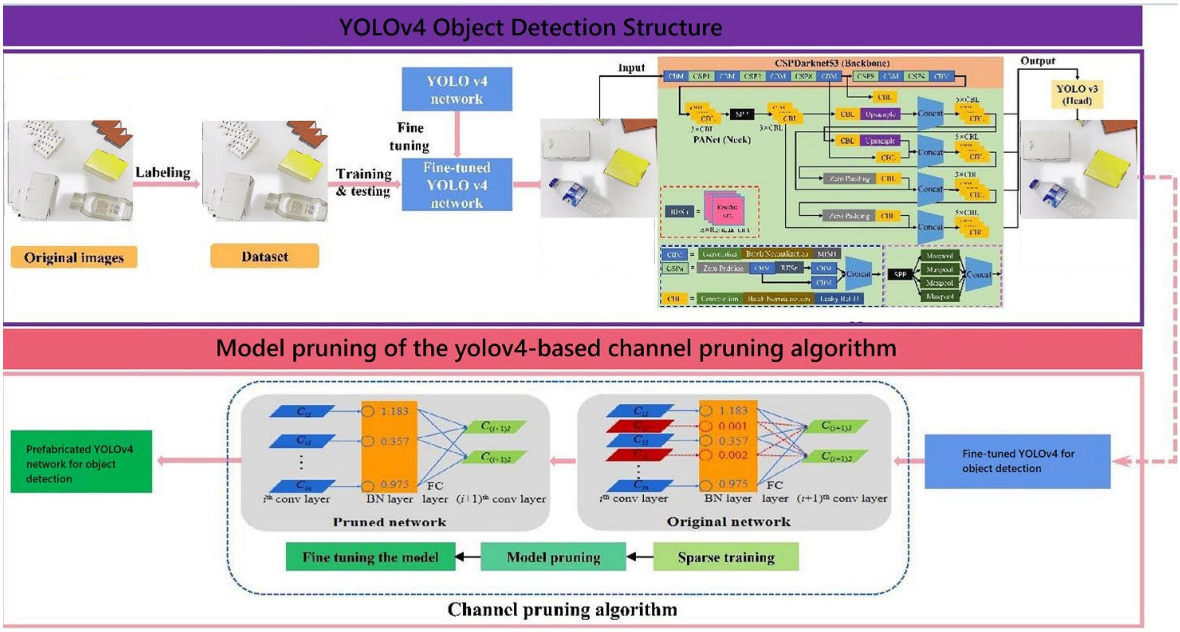

Figure 2: YOLOv4 Structure

2.2 3D Position Measurement

To separate the detected waste using a manipulator, the location of the waste must be known in 3 dimensions. In the study, the Intel RealSense D415 camera, which is a 3D camera, was used to detect the location of the waste. The camera used is shown in figure 4.

Figure 4: Intel RealSense 3D Camera

The transition from the 2D coordinate system to the 3D coordinate system is shown in Figure 5. In the picture shown in Figure 2, the coordinates of the 2-dimensional picture are given as (X_i, Y_i). Then, from the (X_i, Y_i) coordinates, the 3D coordinate (X_s, Y_s, Z_s) system was switched. X_s and Y_s positions are calculated in equation 1, Z_s position is calculated directly by the camera. In Equation 1, (c_x, c_Y) represents the optical center of the camera and (f_x, f_y) represents the focal length of the camera. The c_x, c_Y, f_x, f_y are constant parameters and these all adjusted by camera calibration.

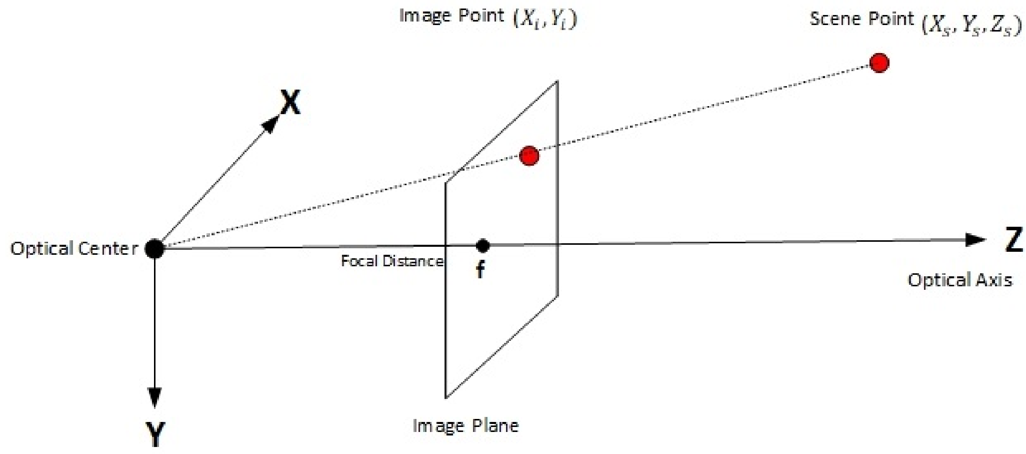

Figure 5: Transition from picture coordinate system to camera coordinate system.

$$X_s = Z_s * \frac{X_i - c_x}{f_x}$$
$$Y_s = Z_s * \frac{Y_i - c_y}{f_y} \quad (1)$$

2.3 Data Set

Data size is a big factor in deep learning models. The more data, the better the model performs. For this reason, the data set needs to be enlarged. The created data set is deformed such as adding blur and noise with image processing techniques, thus providing a better performance of the model even under difficult conditions. In addition, horizontal mirroring and vertical mirroring were applied. The classes and dimensions of the data set used are shown in Table 1.

Table 1: *Size of Classes in the Dataset.*

Class	Data	Size	Total
Paper	Train	552	678
	Test	126	
Metal	Train	332	396
	Test	63	
Glass	Train	378	498
	Test	120	
Plastic	Train	324	402
	Test	78	

All data to be trained are divided into two sets, “training” and “test” data, as shown in Table 2. With these allocated data, the training was started and the training results were explained in detail in Chapter 4.

Table 2: *Proportions of "Training" and "Testing" Datasets to be Included in Training.*

Data	Size (%)
Train	1.580 (%80)
Test	394 (%20)

3. Real-Time Waste Detection

In the study, real-time waste detection was performed after data set training and location detection. The experimental setup to detect waste is shown in Figure 6.

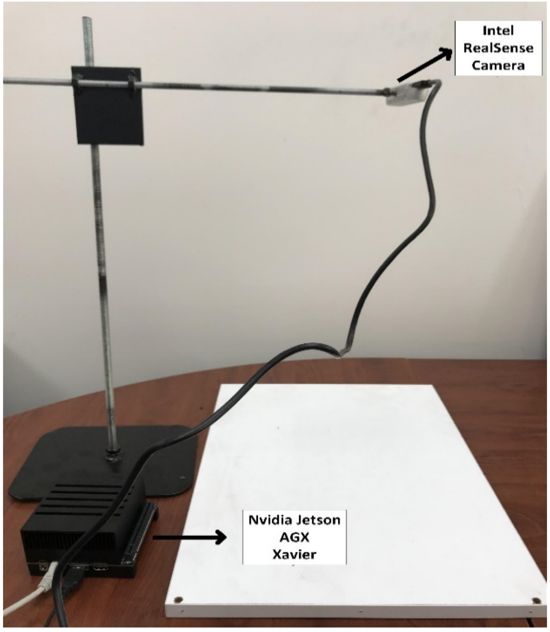

Figure 6: The experimental setup

In the experimental setup, 4 different wastes, metal, paper, glass, and plastic, were detected by using pyrealsense2, OpenCV libraries, and YOLOv4 algorithm on the images taken from the camera in the python environment. The study was tested on the Nvidia Jetson AGX Xavier card, the center and 3D position of the waste were calculated for the decomposition process. Real-time waste detection with images taken from the camera is shown in Figure 7.

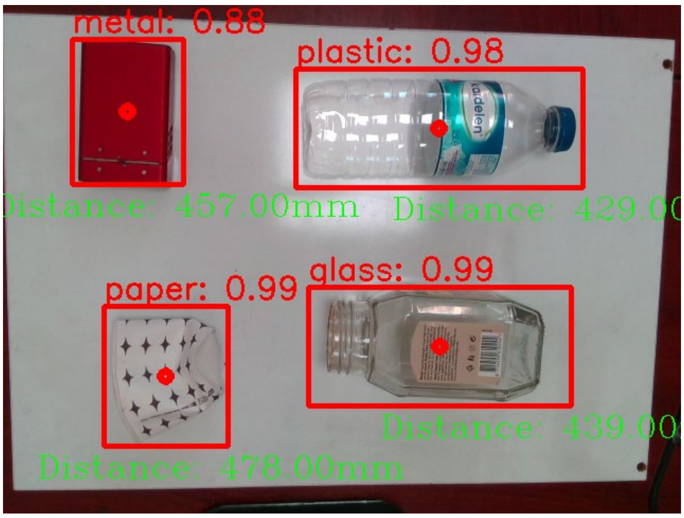

Figure 7: Real-Time Waste Detection

4 Results and Discussion

Within the scope of the study, YOLOv4 and YOLOv4-Tiny algorithms were used and the best result was obtained from the YOLOv4 model. In this model, images with a resolution of 416 x 416 were used as input. The training was completed in 1,000 epochs. Accuracy, verification accuracy, loss, and verification loss results obtained as a result of the training are shown in Table 3.

Table 3: *Training Results*

Model	Accuracy	Verification Accuracy	Loss	Verification Loss
YOLOv4-Tiny	0.9265	0.9296	0.7984	0.7565
YOLOv4	0.9571	0.9587	0.7798	0.7307

Confusion matrix, recall, precision, f-1 score, and intersect over union (IoU) were used as evaluation criteria.

TP (True Positive): These are the cases where the actual value is 1 as well as the predicted value.

TF (True False): These are the cases where the actual value is 0 and the predicted value is not 0.

Accuracy is found by taking the arithmetic mean precision (mAP) of the average precision (AP) values of the classes. Equation 3 shows the formulas.

$$AP = \left(\frac{1}{n}\right)\sum_{Recall_i} Precision(Recall_i) = 1$$

$$mAp = \frac{1}{|classes|}\sum_{c \in classes}\frac{\#TP(c)}{\#TP(c) + \#FP(c)} \tag{3}$$

Recall is known as the hit rate and is expected to be high. It is a measure of how much TP or TN the classifier has correctly estimated. Equation 4 shows how the sensitivity is calculated.

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

Precision is a measure of how accurate or incorrect is the prediction from all classes. Like recall, precision is expected to be high. Equation 5 shows how precision is calculated.

$$precision = \frac{TP}{TP + FP} \tag{5}$$

The classifier's performance is measured by the f-1 score. It is the harmonic mean of sensitivity and precision. Equation 6 shows how the f-1 score is calculated.

$$F1 = 2\,x\,\frac{Recall\ \ x\,Precision}{Recall\ + Precision} \tag{6}$$

The IoU is used for each bounding box to measure the overlap between the predicted bounding box and the actual bounding box. Figure 8 shows the structure.

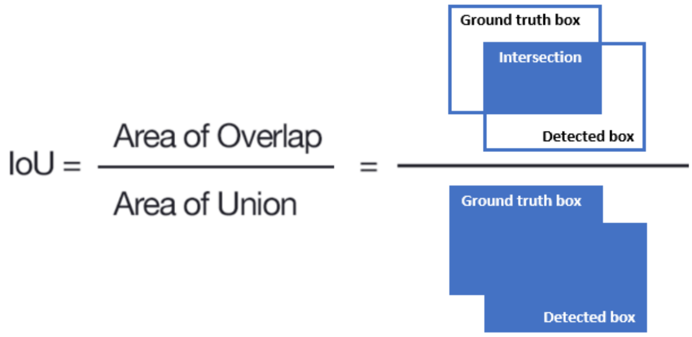

Figure 8: IoU Structure

The results of the evaluation are shown in Table 4, the confusion matrix of the model. Table 5 shows the classification results of the trained models.

Table 4: *Confusion Matrix*

Models	Classes	TP	TF	AP(%)
YOLOv4-Tiny	Paper	547	131	84
	Metal	276	120	62
	Glass	397	101	92
	Plastic	305	97	90
YOLOv4	Paper	558	120	87
	Metal	294	102	68
	Glass	412	86	96
	Plastic	339	63	95

Table 5: *Classification Result of Trained Model*

Model	Sensitivity	Precision	F1	IoU(%)	mAP(%)
YOLOv4-Tiny	0.91	0.63	0.74	74.56	81.56
YOLOv4	0.90	0.91	0.92	78.36	90.23

According to the study's findings, the YOLOv4 algorithm was successful in detecting 4 different types of waste. In future studies, it is planned to create a prototype mechanism to separate the wastes coming from the real conveyor line with a robotic manipulator.

5. Declarations

5.1 Funding source

This study was financed by Sakarya University of Applied Sciences BAP coordinator with project number 022-2022.

5.2 Authors' Contributions

Kenan Erin : Design and implementation of the research, Analysis of the result, Writing of the manuscript.
Bünyamin Bingöl: Design and implementation of the research, Analysis of the result, Writing of the manuscript.
Barış Boru: Design and implementation of the research, Analysis of the result, Writing of the manuscript.

References

[1] Tatzer, P., Wolf, M., Panner, T. (2005). Industrial application for inline material sorting using hyperspectral imaging in the NIR range. *Real-Time Imaging*, 11(2), 99-107.

[2] Scavino, E., Wahab, D. A., Hussain, A., Basri, H., Mustafa, M. M. (2009). Application of automated image analysis to the identification and extraction of recyclable plastic bottles. *Journal of Zhejiang University SCIENCE A*, 10(6), 794-799.

[3] Özkan, K., Ergin, S., Işık, Ş., & Işıklı, İ. (2015). A new classification scheme of plastic wastes based upon recycling labels. *Waste Management*, 35, 29-35.

[4] Meng, S., & Chu, W. T. (2020, February). A study of garbage classification with convolutional neural networks. In *2020 Indo–Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)* (pp. 152-157). IEEE.

[5] Cao, L., & Xiang, W. (2020, June). Application of convolutional neural network based on transfer learning for garbage classification. In *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)* (pp. 1032-1036). IEEE.

[6] Liu, J., Balatti, P., Ellis, K., Hadjivelichkov, D., Stoyanov, D., Ajoudani, A., & Kanoulas, D. (2021, July). Garbage collection and sorting with a mobile manipulator using deep learning and whole-body control. In *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)* (pp. 408-414). IEEE.

[7] S. Shinde, A. Kothari, and V. Gupta, "YOLO based Human Action Recognition and Localization," in Procedia Computer Science, Jan. 2018, vol. 133, pp. 831–838, doi: 10.1016/j.procs.2018.07.112.

[8] Hendry and R. C. Chen, "Automatic License Plate Recognition via sliding-window darknet-YOLO deep learning," Image Vis. Comput., vol. 87, pp. 47–56, Jul. 2019, doi: 10.1016/j.imavis.2019.04.007