

YOLO-BASED SIZE ESTIMATION WITH REFERENCE OBJECTS TO ENHANCE CLASSIFICATION ACCURACY OF WASTE HOME APPLIANCES

YE-EUN PARK¹, SEUNG-JAE YOO¹, YONG-JUN IM^{1,3}, JUNSEONG KIM^{1,3}
AND TAI-WOO CHANG^{1,2,*}

¹Department of Industrial and Systems Engineering

²Intelligence and Manufacturing Research Center

Kyonggi University

154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon, Gyeonggi 16227, Korea

{ yeeun030916; sms7112 }@kgu.ac.kr; *Corresponding author: keenbee@kgu.ac.kr

³Research Center, East-North Resource Recycling Center, Inc.

27 Dongtancheomdansaneop 1-ro, Hwaseong, Gyeonggi 18469, Korea

{ imyj; jjichum }@enrc.kr

Received May 2025; accepted August 2025

ABSTRACT. *With the rapid increase in e-waste, the need for automated classification technologies to enhance the efficiency and accuracy of waste home appliance collection is becoming increasingly important. This study proposes a post-processing algorithm designed to correct classification errors in object detection models, particularly for items – such as refrigerators – whose classification depends on physical size. By incorporating a reference object, the proposed method utilizes the YOLO11 model to estimate the size of refrigerator dimensions, successfully reclassifying 95 out of 125 previously misclassified cases. The algorithm operates as a post-processing step following YOLO-based detection. When a reference object is identified within an image, the algorithm estimates the actual size of the detected appliance and corrects its class label accordingly. The results demonstrate that size-based classification errors can be effectively mitigated using this method, highlighting its strong potential for practical application in e-waste collection environments. This algorithm could be further enhanced by standardizing the positioning of reference objects and incorporating distance estimation, ultimately contributing to the development of a more precise and fully automated classification system.*

Keywords: E-waste collection system, YOLO, Object detection, Reference object placement, Size estimation

1. Introduction. As the amount of e-waste continues to grow, efficient collection and recycling systems are becoming a major challenge globally. In Korea, the amount of e-waste collected is reported to have increased from approximately 102,185 tons in 2022 to 115,280 tons in 2023, as free door-to-door collection services operated through Internet or phone-based reservation systems are gradually expanding [1]. At the collection sites, accurate verification and recording of the collected items is essential, as the items and quantities determine the allowance paid to the collection managers. As a result, there is a need for an advanced system that can automatically verify the evidence at the collection site to cope with the increasing collection volume.

Recently, attempts have been made to apply artificial intelligence technology from real-time object detectors to the collection verification process to meeting this practical need [2]. Currently, the representative object detection technology being utilized is developed around YOLOv8 models, but object detection algorithms are rapidly evolving, and it is necessary to examine the possibility of performance improvement through comparative

verification with the latest models. In particular, for items that are classified based on their actual size, such as refrigerators, existing models recognize only the relative size in the image, which limits accurate classification. In order to reduce the size classification error of refrigerator items by applying the YOLO11 model, this study proposes a size estimation post-processing algorithm that utilizes a reference object to estimate the actual width and redefines the class based on it. The algorithm is methodologically significant in that it improves actual size-based classification accuracy without altering the detection stage, and its practical applicability has been validated through quantitative evaluation using real-world collection site data.

The remainder of this paper is organized as follows. In Section 2, we review the conceptual background of the study and related previous studies. We describe the specific design and implementation process of the proposed algorithm in Section 3. Section 4 presents the experimental setup and results, and finally, Section 5 discusses the conclusions as well as the applicability and limitations of the system.

2. Conceptual Background and Prior Research. In addition to traditional computer vision techniques, deep learning-based object detection techniques are increasingly being used in automated waste sorting and management. Table 1 summarizes major research studies that have utilized artificial intelligence technology in the field of waste management.

TABLE 1. Research on the use of AI in waste management

Subject	Methodology	Ref.
Size estimation	OpenCV-based contour alignment and distance ratio calculation	[3]
	Reference object-based size calculation	[4]
	Webcam-based contour refinement system application	[5]
3D positioning	MonoFlex-based 3D position-size prediction	[6]
	Vanishing point-based perspective correction experiment	[7]
Household waste	YOLOv8-CBAM-based mAP performance experiment	[8]
Electronic waste	YOLOv8-based object detection and active learning	[9]

Rosebrock utilized OpenCV to align contours in an image, extract bounding box vertices, and calculate the distance-per-pixel ratio using the actual length of a reference object (coin) to realize high-accuracy size measurement [3]. Dhikhi et al. proposed a method to automatically calculate the size by selecting an object whose actual size is known in the image as a reference object [4], and Deshmukh et al. developed a system to measure the actual size of an object through real-time image capture via a webcam and automatic contour refinement combining Gaussian blur and erosion-dilation operations [5].

There is also active research on estimating 3D position and size information based on 2D bounding boxes. Zhang et al. extracted feature maps with a CNN (Convolutional Neural Network)-based backbone network and used heat maps and regression heads to predict the 3D center position, orientation, and size of objects [6], while Kim and Kim performed perspective correction through vanishing point calculation and inferred the 3D position and size of real-world vehicles from 2D boxes [7].

Arishi applied five data augmentation techniques and two attention mechanisms to the YOLOv8 model, achieving an mAP of 89.5 on a custom waste dataset comprising 3775 images across 17 classes, outperforming existing two-stage detectors [8].

Research is also being conducted in Korea to apply this technology in practical, real-world scenarios. For example, Im et al. proposed a framework combining YOLO model

with active learning to enhance the efficiency of e-waste collection process, taking an empirical approach to improve the actual collection management system [9].

Previous studies have generally focused on determining the presence of waste or improving the accuracy of item-specific detection, and limited attention has been paid to classification methods that take account of the physical size of objects in an image. In particular, for items such as refrigerators, where the classification criteria vary depending on the actual size even if they are similar in appearance, existing detection methods based on relative size alone have the potential for class misclassification.

3. Research Methodology.

3.1. Background and overview of the algorithm. Existing YOLO models perform class classification based solely on the relative size of detected objects, which prevents them from reflecting the actual size of items with detailed size criteria, such as refrigerators. The proposed method compares the relative size of detected objects in YOLOv11 results with a reference object to estimate their actual width and redefine the class name accordingly, thereby improving the accuracy of refrigerator classification based on size criteria.

3.2. Dataset configuration and detection models used. We used a total of 20,380 images and 24,373 labeled object instances collected at a free collection site in 2023 to train the detection model. We trained the YOLO11 model on this dataset with batch size 16 and epoch 300. The training was performed on an Intel Core i9-13900K CPU, NVIDIA RTX 4090 GPU, and 128GB of RAM. The training resulted in a performance of precision 0.943, recall 0.881, mAP@0.5 0.941, and mAP@0.5:0.95 0.8972.

Evaluation results revealed that the model's accuracy in classifying refrigerators by physical size was inconsistent, with some misclassifications observed for certain size categories. To address this issue, we introduced a post-processing step that estimates the size category of each detected refrigerator using the bounding box output from the trained YOLOv8 model. The size estimation experiment was performed on 125 instances in 121 images of refrigerators, where the predicted and actual values did not match.

3.3. Comparison of detection performance. The accuracy of the detection of the reference object is a prerequisite for the size estimation. Therefore, an experiment on the color setting was conducted first. In the experiment, three colors (red, blue, and purple) were used as the basis, and the performance after detection was compared after randomly inserting the reference objects of each color in the refrigerator. The results are shown in Table 2.

TABLE 2. Experimental results of YOLO detection by reference object color

Color	Total number of detections	Number of false positives	Correct recognition rate (%)
Red	119	22	77.60
Blue	108	34	58.73
Purple	125	1	99.20

The purple reference object has the highest detection accuracy of 99.2%, and the other two colors have relatively low recognition rates because they are likely to be similar to the color of the refrigerator body or the colors contained in the background noise. Based on the results, we set purple as the final target color.

The reference objects used in this experiment were randomly inserted into the image to simulate the actual field application environment. Its size was automatically scaled to be proportional to the relative size of the refrigerator bounding box. The position was

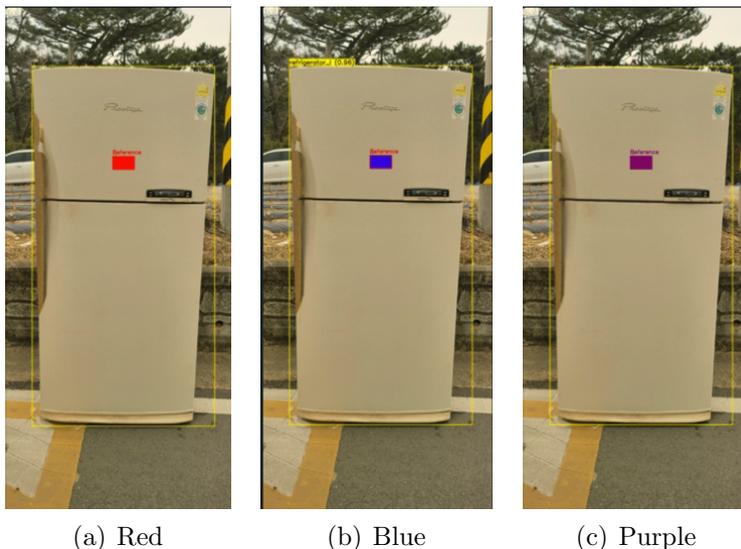


FIGURE 1. Example of detection after attaching a reference object

set to be fixed at the top center of the bounding box. An example of a reference object insertion is shown in Figure 1.

3.4. Size estimation logic and reference object conditions. The proposed algorithm is selectively applied only to the classes that need to be classified. The target of the algorithm is defined as the following subset of the entire class set C :

$$C_{\text{target}} = \{\text{refrigerator_xs}, \text{refrigerator_s}, \text{refrigerator_m}, \text{refrigerator_l}, \\ \text{refrigerator_xl1}, \text{refrigerator_xl2}\}$$

For objects in the set, size estimation is performed after determining whether the reference object exists, and the class name is redefined based on the estimated actual width value. If the reference object does not exist, the original detection result is retained. The algorithm consists of the following sequential steps.

- Step 1. Search for a reference object inside the bounding box, which is defined as a rectangular shape with an aspect ratio of at least 0.8 and no more than 2.5 and a color in the purple range.
- Step 2. Measure the width of the selected reference object and the width of the bounding box and calculate the ratio between them. The width of the reference object is fixed at 11 cm, so the calculated ratio value is multiplied by the actual measurement value to estimate the actual width of the refrigerator.
- Step 3. Compare the estimated width value with the predefined width bins for each class to redefine the class name.

4. Reference Object-Based Size Estimation Results. For re-classification, we set a total of six classification criteria from refrigerator_xs to refrigerator_xl2 as shown in Table 3 based on ground truth information contained in the actual trained data. Each criterion is categorized based on the front width (cm) of the refrigerator.

In this experiment, 95 out of 125 instances with detection errors were correctly classified using the YOLO model. Figure 2 shows the refrigerator images, which were redefined correctly after size estimation.

Confusion matrix analysis showed that confusion between refrigerator_l and refrigerator_xl1 and between refrigerator_xl1 and refrigerator_xl2 occurred frequently, which is interpreted as a result of the proximity of the class bins and the fact that the actual sizes of the refrigerators were concentrated near the boundary values. Some errors also

TABLE 3. Refrigerators are categorized by class.

Class name	Classification criteria	Description
refrigerator_xs	$\text{width} < 45$	Compact storage-purpose refrigerators (mini or auxiliary fridges for small items)
refrigerator_s	$45 \leq \text{width} < 65$	Standard refrigerators for studio apartments or small households (typical single-door or small two-door units)
refrigerator_m	$65 \leq \text{width} < 80$	Two-door household refrigerators and commercial beverage refrigerators (includes typical two-door home fridges and dual-door beverage coolers for business use)
refrigerator_l	$80 \leq \text{width} < 125$	Side-by-side or built-in style household refrigerators (large capacity fridges with double doors or integrated cabinetry design)
refrigerator_xl1	$125 \leq \text{width} < 155$	Four-door large commercial refrigerators (high-capacity, four-door units commonly used in commercial settings)
refrigerator_xl2	$\text{width} \geq 155$	Ultra-large and commercial-use refrigerators (extra-large units or specialty commercial refrigerators beyond standard sizes)



FIGURE 2. Refrigerator image redefined correctly after size estimation

occurred when the reference object was not exactly in the center or front of the image, or when the angle of the shot distorted the ratio calculation.

However, the reference objects were reliably detected in all the experimental images and were correctly matched with each refrigerator even in a multi-object environment, confirming that the size estimation algorithm performed well. The reference objects were judged as valid objects only if they satisfied the set detection conditions (color, aspect ratio, and location centrality), thus ensuring their reliability.

In conclusion, the reference object-based size estimation method was demonstrated to be able to effectively correct size-based classification errors by combining YOLO detection results with ground truth-based correction, and its potential as an applicable post-processing technique in automated collection systems was confirmed.

5. Conclusion. This study proposes an algorithmic approach to improve the accuracy of object detection-based item recognition by utilizing image data collected at a waste

home appliance collection site. In particular, we focus on refrigerator items, which have a fine-grained classification system based on the width of the front, and pay attention to the size misclassification problem that may occur in existing object detection models.

In order to establish classification criteria based on the actual size of refrigerators, we set up six criteria of size classification ranging from extra-small (refrigerator_xs) to extra-large (refrigerator_xl2) based on ground truth values contained in the training data (see Table 3). Based on the above criteria, a post-processing size estimation algorithm utilizing ground truth was applied to the YOLO detection results. The proposed algorithm works by comparing the pixel widths of the detected refrigerator and the reference object to calculate the relative scale factor, which is then multiplied by the actual size of the reference object (11 cm) to estimate the actual width.

In our experiments, 95 out of 125 misclassification cases (76%) were reclassified to the correct class bin, with high consistency and precision in images where the reference object was reliably detected. This is confirmed by the fact that out of 1181 refrigerator instances in the entire test dataset, 1060 were detected using the object detection model alone, while 1155 could be detected using the method. The classification test on 121 real-world images showed that it takes about 40 seconds, which we believe is appropriate for real-world applications. In addition, the algorithm works independently in the post-processing stage without changing the structure of the YOLO model, and can be easily integrated into real-world collection systems.

However, there is still some confusion around the boundary values between refrigerator_l and adjacent classes such as refrigerator_xl1 and refrigerator_xl2, which is analyzed to be caused by factors such as the shooting angle and object alignment. In the future, it is possible to improve the precision of size estimation by normalizing the position of the reference object and applying a distance-based correction factor.

Acknowledgment. This work was supported by the IITP (Institute of Information & Communications Technology Planning & Evaluation)-ICAN (ICT Challenge and Advanced Network of HRD) grant funded by the Korean government (Ministry of Science and ICT).

REFERENCES

- [1] E-Cycle Governance, *2023 E-Cycle Governance Operation Report*, 2024.
- [2] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You Only Look Once: Unified, real-time object detection, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.779-788, 2016.
- [3] A. Rosebrock, Measuring size of objects in an image with OpenCV, *PyImageSearch*, <https://pyimagesearch.com/2016/03/28/measuring-size-of-objects-in-an-image-with-opencv/>, 2016.
- [4] T. Dhikhi, A. N. Suhas, G. R. Reddy and K. C. Vardhan, Measuring size of an object using computer vision, *International Journal of Innovative Technology and Exploring Engineering*, vol.8, no.6S4, pp.661-664, 2019.
- [5] R. Deshmukh, V. Patel and S. Kshirsagar, Real-time object detection and measurement, *Journal Publication of International Research for Engineering and Management*, vol.10, no.4, pp.1-4, 2023.
- [6] J. Zhang, W. Wang, D. Zhang, C. Xu and W. Ouyang, MonoFlex: Monocular 3D object detection through geometry-aware curriculum learning, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.7619-7628, 2021.
- [7] J. Kim and C. Kim, Detection of 3D bounding boxes of vehicles using perspective transformation for traffic surveillance, *Electronics Letters*, vol.55, no.21, pp.1125-1127, 2019.
- [8] A. Arishi, Real-time household waste detection and classification for sustainable recycling: A deep learning approach, *Sustainability*, vol.17, 1902, 2025.
- [9] Y. J. Im, T. W. Chang, S. Park and H. S. Kim, Object detection models and active learning for improvement of e-waste collection management systems in Korea, *Waste Management*, vol.194, pp.379-389, 2025.