JTE HCMUTE
ISSN: 2615-9740

JOURNAL OF TECHNICAL EDUCATION SCIENCE
Ho Chi Minh City University of Technology and Education
Website: https://jte.edu.vn
Email: jte@hcmute.edu.vn

An Intelligent Plastic Waste Detection and Classification System Based on Deep Learning and Delta Robot

Duc Thien Tran[1], Tran Buu Thach Nguyen[2*]

[1]*Ho Chi Minh City University of Technology and Education, Vietnam*

[2]*School of Mechanical and Automotive Engineering, University of Ulsan, Ulsan 44610, South Korea*

[*]Corresponding author. *Email: nguyentranbuuthach2001@gmail.com*

ARTICLE INFO

Received:	18/03/2024
Revised:	16/04/2024
Accepted:	18/06/2024
Published:	28/02/2025

KEYWORDS

Plastic waste classification;

Deep learning;

Transfer learning;

YOLO;

Delta robot.

ABSTRACT

This paper proposes an intelligent plastic waste detection and classification system based on the Deep Learning model and Delta robot. This system includes a Delta robot, a camera, a conveyor, a control cabinet, and a personal computer. The system applies Transfer Learning with the pre-train YOLOv5 model to detect plastic waste in real-time. The best model is selected with the best weight by evaluating the results of the pre-train model to classify different types of plastic waste and determine the positions of the waste by Bounding box. Then, these positions are converted into the Delta robot's coordinate system by the formula obtained from the transformation matrix and the position of the camera. Finally, the computer processes and transports data to control the Delta robot to classify plastic waste in the conveyor. Afterward, a variety of classification experiments with more than 1000 samples in two different lighting conditions were conducted. The results illustrate that the computer vision and deep learning model achieve excellent efficiency with the best-performing case having a Precision of 96% and a Recall of 97%. In conclusion, the experimental results in this paper demonstrate that the proposed intelligent plastic waste detection and classification system delivers high performance both in terms of accuracy and efficiency and has much more potential for further development.

Doi: https://doi.org/10.54644/jte.2025.1555

1. Introduction

In the contemporary era marked by industrialization, modernization, and rapid population growth, the significant increase in both industrial and household waste has become a pressing global concern. Annually, humanity generates an average of 300 million tons of plastic waste. In the year 2021, the world generated an alarming 353 million tons of plastic waste [1]. Regrettably, only approximately 7% underwent recycling, while an overwhelming majority, exceeding 80%, found their way into the oceans and the environment [2]. The significant quantities and diverse composition of waste pose significant challenges, particularly in developing countries. This issue necessitates urgent attention, as it not only impacts the quality of living environments through pollution but also directly impacts human health. Manual collection, sorting, and processing of waste prove to be prohibitively expensive, time-consuming endeavors. Moreover, individuals involved in these tasks face health risks due to the elevated bacterial content inherent in waste materials. Recognizing these challenges, major nations are progressively incorporating automation into industrial processes and daily life. Automated systems, production lines, and robotics are increasingly being deployed, offering the advantage of executing tasks at significantly accelerated rates. Crucially, automation has the potential to take over dangerous tasks, minimizing risks and enhancing workplace safety. It also offers a solution to the challenges associated with waste classification. Numerous approaches and robotic systems have been suggested, demonstrating commendable performance. However, the high cost and complexity of these systems often make installation and maintenance challenging. Moreover, existing waste classification algorithms designed for personal computers fall short of meeting practical needs. Consequently, a waste classification system that combines accuracy, speed, efficiency, affordability, and easy installation has not yet been effectively implemented in the industry.

This research introduces an Intelligent plastic waste detection and classification system, addressing cost challenges while maintaining high levels of accuracy and efficiency. Our approach involves integrating the Delta robot with a plastic waste classification algorithm based on deep learning models. The Delta robot stands out for its simple and compact design, offering exceptional speed and accuracy, making it a widely adopted choice in sorting applications. To control the Delta robot, the Arduino controller—a popular choice is utilized for its compact size and exceptional functionality [3]. The Arduino is programmed to simultaneously control the robot's stepper motors and the pneumatic mechanism which is responsible for object suction and release. Object detection and classification algorithms can be complex and resource intensive. Therefore, the transfer learning method is applied, training the YOLO model based on CNN. YOLO is chosen for its speed advantage over other models, and its open-source nature allows us to train it with our collected dataset. This research focuses on training the model to classify two types of plastic waste: bottle and can. Successful training is followed by a series of tests conducted under various lighting and environmental conditions, demonstrating the system's impressive accuracy. Additionally, precise camera calibration plays a pivotal role in the object position detection process. Through a series of transformations, the object position is obtained in the Robot coordinate system, enabling precise control for the classification process.

Our paper is structured as follows: Section 2 provides an exploration of the YOLO network model and its applications in relevant contexts. In Section 3, we outline in detail the construction and control of our system, offering a comprehensive overview of the employed devices and the process of training the YOLOv5 model for the classification of plastic waste. Section 4 includes a series of experiments conducted under various lighting conditions to verify the accuracy of the detection and classification model and the overall efficacy of the system. Finally, Section 5 is the conclusion and the future directions for development.

2. Related work

This section summarizes relevant research on the application of automatic waste identification and classification. YOLO (You Only Look Once) - One of the famous object recognition and classification models with high speed and accuracy based on Convolutional Neural Network was first introduced by Josep et al. 2016 [4]. Another study by Anbang Ye et al. [5] was conducted to evaluate the waste classification model they designed based on this YOLO model. After the training process, this model achieves a correct rate of 70% with a total number of 32 million parameters and a speed of processing 60 Frames Per Second (FPS). Berardina De Carolis et al. [6] developed software that can detect and classify the presence of abandoned waste through the analysis of video streams in real-time. An improved YOLOv3 network model was trained on the dataset collected for this purpose. Another study used the Single Shot MultiBox Detector (SSD) model which also is a deep-learning model based on a convolutional neural network for garbage detection in various complex backgrounds before being transported by a robot arm and conveyor belt simulated by an electronic rotating turntable [7]. In this study, the results show that the system operates stability with a significant speed of 27.8 FPS, an accuracy rate of approximately 87%, and three types of garbage (ring can, bottle, and aluminum foil pack). A new GNet model for garbage classification based on transfer learning and the improved MobileNetV3 model is developed by Bowen et al. [8]. The classification algorithm will be combined with a camera, an infrared sensor, a laser ranging sensor, and the embedded Linux system which is controlled by a Raspberry Pi 4B microcontroller to perform sorting 4 types of recycled garbage. A series of garbage classification experiments on the Huawei Garbage Classification Challenge Cup dataset were conducted. This classification system's prediction accuracy was 92.62% at 0.63s efficiency. Studies focusing on developing waste classification systems have received more attention around the world in recent years, but these systems are often difficult to access because of the complexity of installation as well as the expensive production cost. In Viet Nam, there have been a few studies on this issue [9]-[12], but these studies have stopped at the laboratory level, and have not been applied in industrial factories. Realizing the potential development of the robot industry as well as image processing, this paper proposed a system that integrates both these technologies to research, develop, and handle the problem of waste detection and classification.

3. System description

3.1. Hardware description

The hardware architecture of the system is delineated into four primary modules: Vision, Controller, Robot, and Actuator, as illustrated in Figure 1.

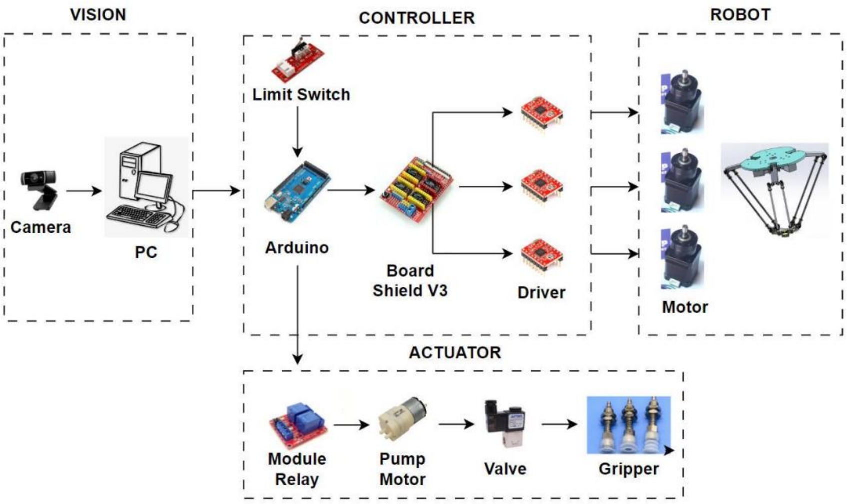

Figure 1. *The overview of the Intelligent Plastic Waste detection and Classification System*

The Vision module is responsible for receiving input information in the form of images from the camera and transmitting it to the initial processing center, i.e., the PC. The YOLO model is utilized for object classification and extraction of object position.

The Controller module includes the Arduino MEGA, which is the master board for the hardware system, CNC shield, and stepper drivers. Upon acquiring the object position, the number of pulses for each motor is calculated and dispatched by Arduino to control the robot's movement, ensuring coordinated operation with other devices cyclically.

The Robot module includes the Delta Robot, consisting of a robot body and three stepper motors. Upon receiving the calculated signals from the Controller module, the Delta Robot moves to approach the object and perform classification.

The Actuator module is integrated with a Relay Module to regulate the vacuum pump and a 2-way valve. The vacuum pump generates suction to hold the object in the robot's end effector. Simultaneously, as the robot reaches its final position in the classification process, the valve releases the air, allowing the object to fall into the appropriate trash can.

This hardware structure facilitates a systematic and efficient workflow, where each device plays a distinct role in the overall operation of the robotic system.

3.2. Object classification

One proposed solution to address the issue of detecting and classifying plastic waste is the utilization of the YOLO model, which stands for You Only Look Once. It's one of the fastest object detection and classification models based on CNN (Convolutional Neural Network) [13]. The Logitech C920 camera, operating at a resolution of 1920x1080p, is employed alongside a conveyor system where plastic waste is placed and passed through the camera's operational area. Subsequently, the YOLO model is employed to precisely determine the coordinates of the plastic waste within the camera's frame before these coordinates are calculated, transformed, and relayed to the Robot coordinate system for further processing.

The YOLO model encompasses various versions and undergoes continuous updates. However, YOLOv5 is selected due to significant enhancements in both accuracy and processing speed compared

JTE
HCMUTE
ISSN: 2615-9740
JOURNAL OF TECHNICAL EDUCATION SCIENCE
Ho Chi Minh City University of Technology and Education
Website: https://jte.edu.vn
Email: jte@hcmute.edu.vn

to its predecessors [14], [15]. In addition, with its open-source code as well as being very easy to deploy for embedded systems, YOLOv5 has become one of the classic and well-known models in the deep learning community. Below is the structure of the YOLOv5 object detection and classification model [16].

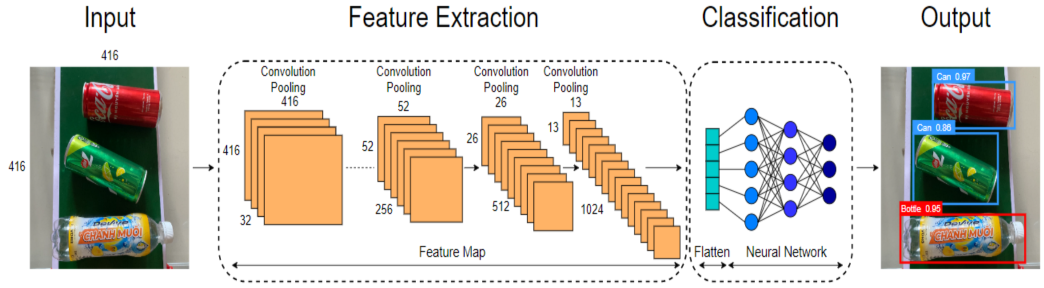

Figure 2. *Yolov5 model structure*

The YOLOv5 model structure is built based on the CNN structure including 2 main blocks. The first block is the feature extraction consisting of many convolutional layers extracting detailed features of the image based on various feature matrices with different sizes. Small feature matrices will predict huge objects and vice versa. In addition, this block also has pooling layers reducing calculation errors and improving model speed. The rest block includes dense layers of neurons based on the extracted features in each pixel cell of the feature matrices to predict three pieces of information: Object position (Bounding Box), object type (Class), and accuracy (Confidence Score). These three pieces of information are the result of the object detection and classification process with the input of an RGB color image.

The subsequent step involves outlining the training process for the YOLOv5 network model, which comprises six main steps [17], [18], as illustrated in Figure 3: Collecting data, Preprocessing, Labeling, Applying augmentation, Dividing datasets, and Training model.

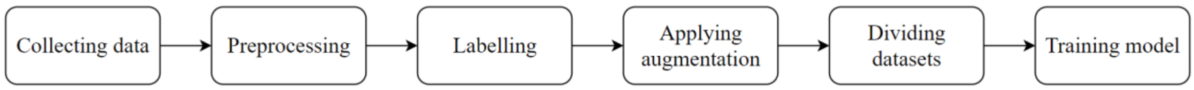

Figure 3. *Steps to train the YOLO model*

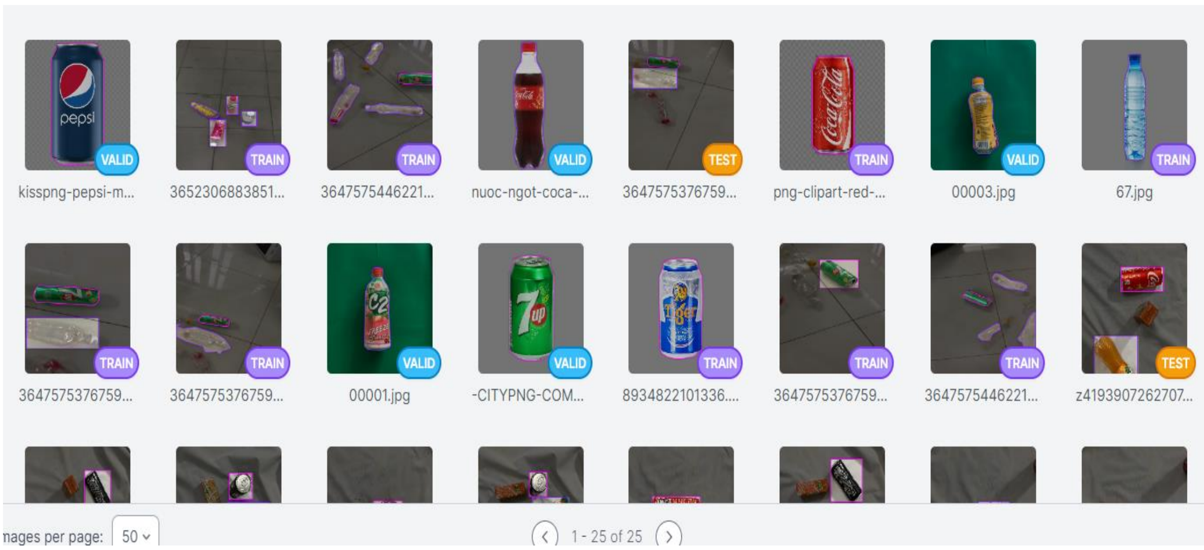

Figure 4. *Data collection process*

The first step involves the creation of an input dataset essential for training the YOLOv5 model. Training images are captured using smartphones with a resolution of 1920x2560. These images are collected under lighting and environmental conditions that closely emulate the operational conditions in which the model is expected to function. This approach aims to enhance the model's accuracy under realistic conditions. The project's focus is on training the model to detect and classify two types of plastic waste: bottles and cans. Consequently, the dataset exclusively includes images featuring these types of plastic waste, with an exclusion of noisy images. Furthermore, to enhance the model's versatility, images of plastic waste are collected from various angles and poses, facilitating the model's proficiency in detecting and classifying plastic waste from diverse cases. The data collection process is shown below in Figure 4.

Upon the collection of approximately 1000 images for the input dataset, a preprocessing step is performed. This step involves resizing all images to a uniform dimension of 416x416, as the YOLOv5 network model exclusively supports processing with this input image size.

Following the preprocessing step, all the plastic waste in the input images is labeled by using the Boulding box labeling method. This pivotal yet time-consuming process entails precisely localizing and annotating each plastic waste present in the images.

Figure 5. *Training process*

In the fourth step, several augmentation tools such as flip, crop, rotation, saturation, and brightness are applied to all images in the input dataset to enrich the quantity and quality of the dataset. Following the augmentation process, more than 3,000 images are utilized for the training process.

In the penultimate step, all the images are divided into 3 sets: the training set (accounts for 70 percent) is utilized for training the YOLO model, the validation set (accounts for 20 percent) estimates and picks the best suitable and accurate model for the network, and the test set (accounts for 10 percent) checks and evaluates the accuracy and the error of the model.

Finally, the training process is conducted on the Google Colab platform [19], which is a product of Google Research, it allows Python to be performed on the cloud platform, especially suitable for data analysis, machine learning, and education. The model experienced 150 training cycles with 16 batches and then achieved around 96% of Precision and 97% of Recall.

3.3. Object localization

Object Localization involves the precise determination of an object's position within the Robot coordinate system, achieved through transformative processes based on the camera working area and Robot coordinate system, which is shown in Figure 6.

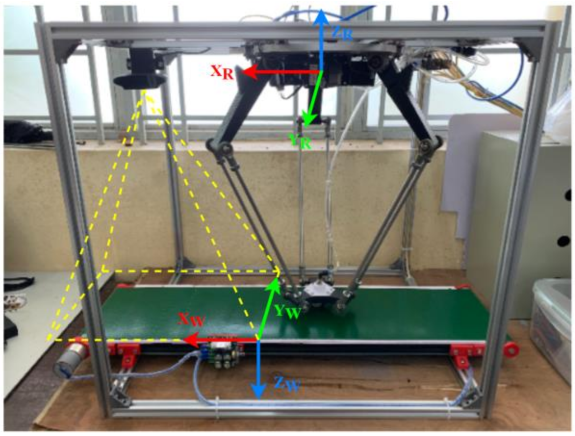

Figure 6. *Camera working area and Robot coordinate system*

Subsequently, these coordinates serve as the foundation for controlling the robot in approaching the objects. This process is divided into two sequential steps, as illustrated in Figure 7.

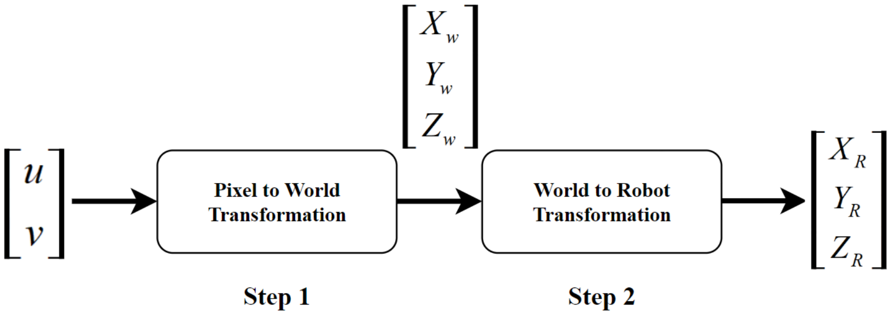

Figure 7. *Object localization steps*

Where:

u, v : coordinates of the objects in the Pixel coordinate system.

X_W, Y_W, Z_W : coordinates of the objects in the World coordinate system.

X_R, Y_R, Z_R : coordinates of the objects in the Robot coordinate system.

Upon successful detection, classification, and determination of the object's position by the YOLOv5 model, a preprocessing step is performed to calculate the object's centroid position as the center of the bounding box within the camera frame, which has dimensions of 640x480 pixels.

Figure 8. *Determine the centroid of the object*

In the initial step, the coordinates of the objects within the pre-defined and established World coordinate system, as shown in Figure 6, are obtained [20]. Because the object to be classified is always located on the conveyor belt, the coordinates of the object along the Z-axis of the World coordinate system are always set to 0 ($Z_W = 0$). The correlation between the two coordinate systems is leveraged to derive a conversion formula:

$$\begin{bmatrix} X_W \\ Y_W \end{bmatrix} = \frac{S}{m} \times \begin{bmatrix} u \\ v \end{bmatrix} \tag{1}$$

Where:

S : Actual distance (mm).

m : Pixel density with S distance.

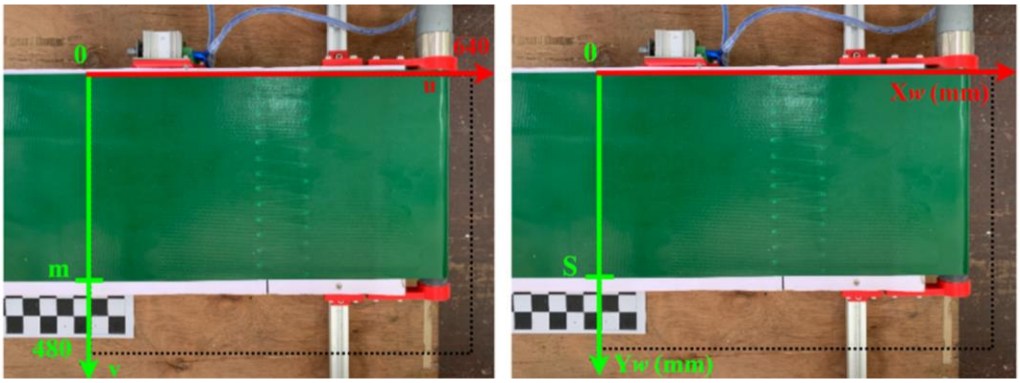

Figure 9. *Correlation between Pixel and World coordinate system*

In the second step, an additional transformation is executed to convert the object's coordinates from the World coordinate system to the Robot coordinate system before controlling the robot to move and perform the classification process.

The translation matrix [21] is determined based on the positions of the two established coordinate systems. Subsequently, the conversion formula facilitating the transformation of points from the World coordinate system to the Robot coordinate system is derived:

$$\begin{bmatrix} X_R \\ Y_R \\ Z_R \\ 1 \end{bmatrix} = {}_W^R T \times \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_X \\ R_{21} & R_{22} & R_{23} & T_Y \\ R_{31} & R_{32} & R_{33} & T_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (2)$$

Where ${}_W^R T$ is a transformation matrix, R is rotation matrix, and T is translation matrix.

The object for classification is situated on the conveyor. Therefore, fixed positions are established for the end effector's coordinate of the robot along the Z-axis throughout the steps controlling the robot's operations.

4. Experimental results

4.1. Kinematic

To evaluate the precision of the kinematic conversion process, the robot is controlled to a random position, with both the end effector's coordinates and the theta angles of the robot recorded at this position. In the next stage, new theta angles are calculated based on the inverse kinematic transformation formulas, and the recorded end effector's coordinate. Table 1 is devised for the purpose of contrasting and comparing the disparities between the recorded theta angles and calculated theta angles.

Table 1. *Comparison between the recorded theta angles and calculated theta angles*

P_x, P_y, P_z (mm)	Recorded angle (°)			Calculated angle (°)		
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
(53.77, -84.12, -420.41)	30	45	60	30.3	45.31	60.31
(39.81, -186.35, -385.16)	10	60	70	10.21	60.26	70.25
(-48.53, -60.25, -496.25)	65	90	75	66.11	91.91	76.36
(-70.83, 151.39, -410.19)	75	50	25	75.51	50.53	25.46

After a comprehensive comparison and evaluation, it was discovered that the error between the recorded theta angles and the calculated theta angles was negligible (less than 2°). This indicates that the kinematic calculation process was executed accurately.

4.2. YOLO training

As mentioned before, the YOLOv5 model was trained with more than 1000 input images in the Google Colab platform, which allows for the customization of various parameters.

To demonstrate the effectiveness of the YOLOv5 model, experiments were conducted on the classification results using 400 samples, comprising 200 bottles and 200 cans. The confusion matrix [22] which is the most basic and efficient method to evaluate the accuracy of a classification model is employed. The three parameters, precision, recall, and F1-Score are derived from the confusion matrix. Precision serves as an indicator of accuracy, evaluating the correctness of positive predictions. Recall, also known as the recall rate or sensitivity, is calculated to measure the model's ability to capture and correctly identify all positive predictions. These metrics play a crucial role in assessing the performance of a detection and classification model, providing insights into its accuracy and recall capabilities. Higher values of these metrics indicate better accuracy of the model. However, these two metrics often do not increase or decrease together, so a third metric is needed, which is the F1-Score, to provide an overall evaluation and balance between the two aforementioned metrics.

$$precision = \frac{TP}{TP + FP} \tag{3}$$

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

$$F1 - Score = \frac{2 \times (precision \times recall)}{precision + recall} \tag{5}$$

Where:

TP (True Positive): Number of predictions where the model correctly predicts the positive class as positive. In this case, TP is the number of correctly predicted bottles as bottles.

TN (True Negative): Number of predictions where the model correctly predicts the negative class as negative. In this case, TN is the number of correctly predicted cans as cans.

FP (False Positive): Number of predictions where the model incorrectly predicts the negative class as positive. In this case, FP is the number of incorrectly predicted cans as bottles.

FN (False Negative): Number of predictions where the model incorrectly predicts the positive class as negative. In this case, FN is the number of incorrectly predicted bottles as cans.

Table 2. *Confusion matrix for YOLO model with 512 input images*

Predicted \ Actual	Bottle	Can
Bottle	163	68
Can	37	132

Precision = 0.7056, Recall = 0.815, F1-Score = 0.756

Table 3. *Confusion matrix for YOLO model with 1026 input images*

Predicted \ Actual	Bottle	Can
Bottle	190	4
Can	10	196

Precision = 0.9794, Recall = 0.95, F1-Score = 0.964

As can be seen from both parameters in the two cases, it becomes evident that the initial results are unsatisfactory. In response to this, the volume of training input images is doubled compared to the original dataset (512 images), resulting in a total of 1026 images. Consequently, there is a significant enhancement in the recognition accuracy, with cans and bottles being classified nearly perfectly.

4.3. Lighting condition

To assess the efficacy of the model in various aspects, an additional experiment is conducted, wherein the system is operated under two different lighting conditions: laboratory lighting (intensity of 200 lux) and low-light (intensity of 20 lux). The outcomes are documented and organized into Table 4 and Table 5 for analysis and comparison.

Table 4. *Testing results in laboratory lighting conditions*

Test No.	Bottle		Can	
	Actual	Predicted	Actual	Predicted
1	50	49	50	48
2	50	48	50	49
3	50	48	50	48
4	50	49	50	48

Precision = 0.9652, Recall = 0.97, F1-Score = 0.967

Table 5. *Testing results in low-light conditions*

Test No.	Bottle		Can	
	Actual	Predicted	Actual	Predicted
1	50	45	50	48
2	50	47	50	46
3	50	47	50	47
4	50	43	50	45

Precision = 0.9286, Recall = 0.91, F1-Score = 0.919

Through a comparison of the confusion matrices, precision, recall and F1-Score parameters were computed in both conditions. The findings reveal that the system performs outstandingly under laboratory lighting conditions. Conversely, in low-light conditions, the model encounters difficulties in object recognition and classification, resulting in instances of missed objects.

5. Conclusion

This research introduces a novel intelligent system for plastic waste detection and classification, which integrates a Delta robot arm with the YOLOv5 deep learning model. This innovative approach offers a cost-effective solution for automating plastic waste sorting processes within industrial plants. A set of experiments has been conducted to assess the system's efficiency, accuracy, and productivity. The results demonstrate the system's exceptional performance under laboratory lighting conditions, yielding a Precision of 0.9652, a Recall of 0.97 and a F1-Score of 0.967. The system exhibits a remarkably high correct recognition rate, which tends to escalate with an increased number of training input images. While occasional instances of confusion persist, the overall system response is highly commendable. Notably, the system demonstrates excellence not only in identifying and classifying two types of waste but also in operating effectively and stably.

Acknowledgments

Thank you for the support from Ho Chi Minh City University Technology and Education, Vietnam.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data is contained within the article.

REFERENCES

[1] L. Lebreton and A. Andrady, "Future scenarios of global plastic waste generation and disposal," *Palgrave Communications*, vol. 5, no. 1, pp. 1-11, 2019.

[2] R. Siddique, J. Khatib, and I. Kaur, "Use of recycled plastic in concrete: A review," *Waste management*, vol. 28, no. 10, pp. 1835-1852, 2008.

[3] P. Z. Murray, "Descriptive geometric kinematic analysis of clavel's "Delta" Robot," *Centre of Intelligent Machines, McGill University, USA*, vol. 1, 2004.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779-788.

[5] A. Ye, B. Pang, Y. Jin, and J. Cui, "A YOLO-based neural network with VAE for intelligent garbage detection and classification," in *Proceedings of the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence*, 2020, pp. 1-7.

[6] B. De Carolis, F. Ladogana, and N. Macchiarulo, "Yolo trashnet: Garbage detection in video streams," in *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, 2020: IEEE, pp. 1-7.

[7] S. H. Lee and C. H. Yeh, "A highly efficient garbage pick-up embedded system based on improved SSD neural network using robotic arms," *Journal of Ambient Intelligence and Smart Environments*, no. Preprint, pp. 1-17, 2022.

[8] B. Fu, S. Li, J. Wei, Q. Li, Q. Wang, and J. Tu, "A novel intelligent garbage classification system based on deep learning and an embedded linux system," *IEEE Access*, vol. 9, pp. 131134-131146, 2021.

[9] K. K. D. Vu, T. N. Huynh, A. A. Thuy, and H. P. Tan, "Deep based learning: ebin sorting system development," *Vietnam Journal of Marine Science and Technology*, vol. 22, no. 1, pp. 85-93, 2022.

[10] A. H. Vo, M. T. Vo, and T. Le, "A novel framework for trash classification using deep transfer learning," *IEEE Access*, vol. 7, pp. 178631-178639, 2019.

[11] T. K. Tran, K. T. Huynh, D. N. Le, M. Arif, and H. M. Dinh, "A Deep Trash Classification Model on Raspberry Pi 4," *Intelligent Automation & Soft Computing*, vol. 35, no. 2, 2023.

[12] M. H. Huynh, P. T. Pham-Hoai, A. K. Tran, and T. D. Nguyen, "Automated waste sorting using convolutional neural network," in *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*, 2020: IEEE, pp. 102-107.

[13] M. M. Yusro, R. Ali, and M. S. Hitam, "Comparison of Faster R-CNN and YOLOv5 for Overlapping Objects Recognition," *Baghdad Science Journal*, vol. 20, no. 3, pp. 0893-0893, 2023.

[14] M. Sozzi, S. Cantalamessa, A. Cogato, A. Kayad, and F. Marinello, "Automatic Bunch Detection in White Grape Varieties Using YOLOv3, YOLOv4, and YOLOv5 Deep Learning Algorithms. Agronomy 2022, 12, 319," ed: s Note: MDPI stays neutral with regard to jurisdictional claims in published …, 2022.

[15] U. Nepal and H. Eslamiat, "Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs," *Sensors*, vol. 22, no. 2, p. 464, 2022.

[16] D. Thuan, "Evolution of Yolo algorithm and Yolov5: The State-of-the-Art object detention algorithm," 2021.

[17] J. N. Jacob Solawetz. "How to Train a YOLOv5 Model On a Custom Dataset." https://blog.roboflow.com/how-to-train-yolov5-on-a-custom-dataset/ (accessed Jul. 18, 2023).

[18] J. Glenn. "YOLOv5 by Ultralytics." github.com. https://github.com/ultralytics/yolov5 (accessed Jul. 18, 2023).

[19] J.Glenn."YOLOv5Tutorial." https://colab.research.google.com/github/ultralytics/yolov5/blob/master/tutorial.ipynb (accessed Jul. 18, 2023).

[20] A. Anwar, "What are Intrinsic and Extrinsic Camera Parameters in Computer Vision," *Towards Data Science*, 2022. [Online]. Available: https://towardsdatascience.com/what-are-intrinsic-and-extrinsic-camera-parameters-in-computer-vision-7071b72fb8ec.

[21] J. J. Craig, *Introduction to Robotics*, Third ed. United States of America: Pearson Education, 2005.

[22] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 233-240.

Tran Duc Thien received the B.S and M.S degrees in the Department of Electrical Engineering, Ho Chi Minh City University of Technology, Vietnam, in 2010, and 2013. In 2020, he received Ph.D. degrees in the School of Mechanical Engineering, University of Ulsan, Korea. Currently, he is a lecture in Ho Chi Minh City University of Technology and Education (Vietnam). His research interests include robotics, fluid power control, nonlinear control, adaptive control, intelligent technique, observer, and networked control system. Email: thientd@hcmute.edu.vn. ORCID: https://orcid.org/0000-0002-6684-0681

Nguyen Tran Buu Thach received the Bachelor's degree in Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology and Education, Vietnam, in 2023. Currently, he is studying at the School of Mechanical Engineering, University of Ulsan, Korea. Email: nguyentranbuuthach2001@gmail.com. ORCID: https://orcid.org/0009-0002-4248-6005