Paper ID:151

Automated Waste Sorting using Deep Learning and Robotic Manipulation: A Comprehensive Approach

Md Tanvir Hossain Taaha[1], Rafi B. Dastagir[1], Md Amin Haque[1], M. Abrar Muhit[1]
[1]Department of Mechanical Engineering, Bangladesh University of Engineering and Technology
E-mail: tanvir.taaha@gmail.com

Abstract

This paper presents an automated waste sorting system that combines deep learning for real-time waste detection with inverse kinematics for precise object handling. A fine-tuned YOLOv9 neural network, trained on a custom dataset, identifies and classifies waste based on recyclability and biodegradability. The detected coordinates are processed by the RRT-Connect algorithm, integrated with the ROS (Robot Operating System) MoveIt package, to calculate optimal joint movements for a 4 degrees-of-freedom (DOF) robotic arm. A prototype, designed using CAD and fabricated via 3D printing demonstrated robust performance, successfully sorting waste into designated categories. This system reduces manual labor, minimizes health hazards associated with improper waste handling, and enhances detection accuracy. The suggested solution shows potential in improving waste management in diverse industrial and environmental applications.

Keywords: Waste Sorting, Robotic Arm, Deep Learning, Inverse Kinematics, Motion Planning

1. Introduction

As human productivity increases every year, so does the pile of waste we - as a species - collectively produce. With the world's population and industrial output growing, the need for effective waste management has become critical. Recycling, a crucial component of waste management, is financially beneficial and essential for environmental sustainability. Proper sorting is the first step in recycling, as mixed waste cannot be treated efficiently, and manual sorting not only fails to meet the need but also exposes workers to significant health risks [1]. A powerful solution to this challenge lies in recent advancements in artificial intelligence (AI), helping robotics streamline processes while replicating anthropoid actions, particularly within risky environments [2]. This progress has proven to be especially beneficial for automating waste detection and sorting processes [3].

Waste detection, the first step of waste sorting, can be actualized in one of two ways: by using specialized sensors [4] or through employing various object detection algorithms with computer vision [5]. Popular models based on Convolutional Neural Network (CNN) have been tried for real-time object detection to identify and classify waste in previous studies [6]. Once waste is detected and classified, it must be separated from the pile and placed into the appropriate container. Several methods are available for this physical separation, including robotic arm-based sorting, eddy current separation, and conveyor systems. Robotic arm-based sorting provides the highest precision, enabling accurate object placement. Extensive research has focused on motion planning by calculating the joint angles needed to reach target positions with specific poses [7]. Inverse Kinematics has been successfully implemented to calculate the joint angles of 3 degrees of freedom (DOF) [8] and 4 DOF [9] robotic arms. Recent studies have explored the integration of CNN-based object detection with robotic manipulation for sorting trash. Huang et al. [10] various CNN models for trash detection and implemented them in a simulated trash sorting environment. Liu et al. [11] tested their ResNet model on a simulated robotic arm, while Ibrahim et al. [9] used older YOLO (You Only Look Once) versions in similar simulations.

While advances have been made in CNN-based waste detection and motion planning for trash sorting in simulated environments, there is a notable gap in studies implementing these technologies in real-world robotic systems and evaluating their performance. Our study addresses this by utilizing the state-of-the-art YOLOv9 [12] model to instantly classify and locate different trash types for fine sorting, followed by motion planning for a 4-DoF robotic arm using the MoveIt package. The system is then implemented in a prototype robotic arm for real-world testing. A comprehensive performance analysis of trash classification and mechanical maneuvers is conducted to provide a thorough, application-focused understanding of AI-driven waste sorting.

2. Materials and Methods

Trash Detection

A dataset of 4,120 images was assembled from an online repository [13] to train a trash detection model for classifying and localizing clumped waste. Each 640x640 pixel image contained multiple overlapping trash objects with annotated bounding boxes and class labels. Three augmented versions of each image were created with a 50% chance of horizontal flip, random rotation between -45° and +45°, brightness adjustment from -20% to 0%, and Gaussian blur (0-2 pixels). Bounding boxes were randomly adjusted by -32% to +32%. The dataset was split into 3,900 training images, and 160 images each for validation and testing. Figure 1(a) outlines this process.

To meet real-time object detection needs in dynamic environments, the YOLO algorithm was selected for its efficiency and speed in rapid classification tasks. YOLO analyzes an image in a single pass by splitting it into a grid, with a CNN predicting both the bounding boxes and the class probabilities for each grid cell with deeper layers capturing complex patterns and reducing spatial dimensions for better accuracy and efficiency. While retaining bounding boxes with maximum confidence, overlapping boxes are filtered out by Non-Maximum Suppression (NMS). Figure 1(b) illustrates the YOLO workflow [20]. Training was conducted on a desktop with a Ryzen 7 5700X3D CPU, 32GB RAM, and an NVIDIA RTX 3060ti GPU, running Windows 11 (64-bit). The experiments used PyTorch 2.4.0, including a batch size of 16, epochs of 200 and an initial learning rate of 0.01. Pre-trained weights from YOLOV8 (yolov8m.pt, yolov8x.pt) and YOLOv9 (yolov9-m.pt) were used to accelerate training and enhance performance. Evaluation metrics include the inference time, Precision, Recall, mAP50 and mAP50-95. Precision signifies the ratio of predictions that are true positive to the net predicted positives (the sum total of actual true positives and false positives). Recall assesses the capacity of a model to accurately identify positive cases by computing the ratio of predictive true positives to the sum total number of actual positives. Mean Average Precision (mAP) evaluates overall object detection performance (Equation 1). mAP50 is the mAP at a 50% Intersection over Union (IoU) level, considering a particular reading accurate if the estimated bounding box intersects with the ground truth by no less than 50%. mAP50-95 averages mAP over IoU thresholds from 50% to 95% in 5% increments, offering a more detailed assessment of localization precision.

$$mAP = \frac{\sum_{K=1}^{N} PR}{N} \tag{1}$$

where, P, R and N are Precision, Recall and number of samples respectively. Coordinates of the bounding box are transmitted to the motion planning algorithm for inverse kinematics calculations, facilitating precise control of the robotic-arm used in the trash sorting process.

Motion Planning:

Once the trash coordinates are identified in the image, real-world coordinates are computed using camera calibration and pose estimation, a technique known as Hand-Eye calibration, commonly employed in robotic manipulation [14]. Motion planning then determines the optimal end-effector path and position, factoring in obstacles and joint limitations. Rapidly-exploring Random Tree (RRT) algorithm, effective in high-dimensional and dynamic environments, is utilized for multistate path planning, as demonstrated by Gao et al. [15] and Zou et al. [16]. The RRT method initializes a tree at the start configuration and iteratively generates random samples, extending the tree towards them while checking for collisions. Once a valid path is identified or iteration limits are reached, Inverse Kinematics (IK) computes joint angles to achieve the required end-effector poses (Fig. 2a).

Fig. 2b illustrates a general 4-DoF robotic arm, for which the goal point is denoted as cartesian coordinate, *G(x, y, z)*. It is possible to convert the goal point's position into cylindrical coordinates *(θ, ρ, z)*. The robotic arm structure can be made aligned with the *x-z plane* by a base rotation of *θ degree*, as shown in Fig. 2(b). Line *c*

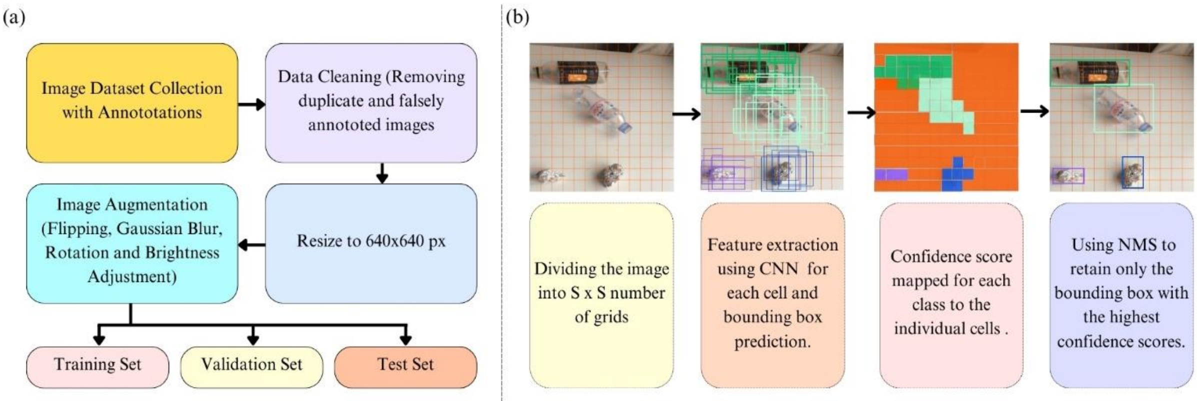

Fig. 1. (a) Schematic representation of the dataset preparation workflow for training the trash detection model. (b) Flowchart depicting the YOLO algorithm's process for detecting and classifying trash.

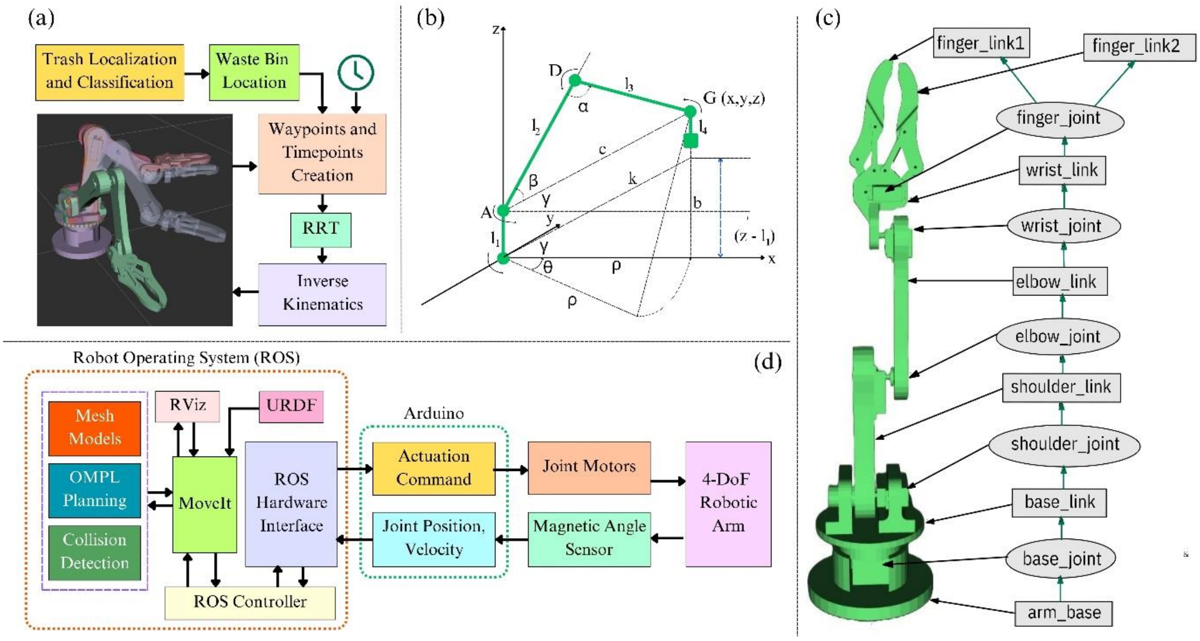

Fig. 2. (a) Workflow of the robotic arm motion planning process. (b) Schematic representation of the geometry for inverse kinematics equations. (c) Schematic of the 4-DOF robotic arm used in the experimental setup. (d) Software stack workflow of the robotic arm module

connects the end-effector position at point *G* to shoulder joint position at point *A*. Line *k* is taken as a parallel line of *c* from the point of the base joint. Equations (2-6) can be used to calculate the joint angles by utilizing the cosine formula. Equation (2) gives projected polar coordinates in *x-y plane* of the goal point. Equation (3) determines the length of line *c*. Equation (4) gives angle, γ between *x-axis* and line *k*, and in triangle ΔAGD the angles α and β are determined by equations (5) and (6).

$$\rho = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right) \tag{2}$$

$$c = \sqrt{(z - l_1)^2 + \rho^2} \tag{3}$$

$$\gamma = \tan^{-1}\left(\frac{z - l_1}{\rho}\right) \tag{4}$$

$$\alpha = \cos^{-1}\left(\frac{l_2^2 + l_3^2 - c^2}{2l_2 l_3}\right) \tag{5}$$

$$\beta = \cos^{-1}\left(\frac{l_2^2 + c^2 - l_3^2}{2l_2 c}\right) \tag{6}$$

The angles of joints are determined by equations (7–10) to bring the end-effector to the goal point. Each joint frame's rotation angle is measured counterclockwise with respect to the preceding link frame. First joint angle θ_1 is exactly equal to the angle θ. The second joint angle θ_2 is the sum of β and δ with an offset of 270 degree. The third joint angle, θ_3 is the angle α, 180 degrees shifted.

$$\theta_1 = \theta \tag{7}$$

$$\theta_2 = 270 + \beta + \gamma \tag{8}$$

$$\theta_3 = 180 + \alpha \tag{9}$$

$$\theta_4 = 90 - (\alpha + \beta + \gamma) \tag{10}$$

where the angles θ_1, θ_2, θ_3 and θ_4 are the required for modifying the joint configuration in order to the end effector be placed at the goal point. the intended position. RRT-Connect, a variation of RRT, is chosen due to its success in most cases, as demonstrated by Kuffner et al. [17]. MoveIt, a ROS package suite, provides convenient access to Open Motion Planning Library, OMPL.

Experimental Setup

A 4-DoF robotic arm was built for real-time trash detection and collection. The arm's structure, made from durable Acrylonitrile butadiene styrene (ABS) plastic, was designed in SolidWorks and 3D printed. Bearings and motor mounts were press-fitted into the joints, and the parts were assembled with metal screws and heat inserts. The 3D design and schematic are shown in Fig. 2c. The robotic arm is powered by four servo motors for pitch, yaw, and claw movements, and a NEMA-23 stepper motor for full base rotation. DS51150 servos control the shoulder and elbow joints, while MG996R servos handle wrist pitch and grip actions. All joints' feedbacks are provided by AS5600 magnetic angle sensors for closed-loop control, except the gripper. An Arduino Mega 2560, interfaced with a laptop running the object detection model, controls the arm. Using ROS, specifically MoveIt, the arm's motion is planned and executed. ROS controllers manage hardware instructions and joint state feedback through a hardware interface node that communicates over Rosserial protocol (Fig. 2d). The microcontroller translates commands into motor signals, calculates position and velocity from sensor data, then relays this information to the hardware interface.

For experimental trials, three trash classes - paper, plastic, and metal - are detected, classified, and dropped to corresponding bins. The bins, labeled paper, plastic, and metal, are placed diversely around the robotic arm to test its capabilities within its workspace. A camera on a tripod provides an overhead view for real-time object detection and classification, simulating an industrial work environment. The trials occur at different times during the day, excluding nighttime, with objects picked from a wooden platform.

3. Results and Discussion

Trash Detection Model Performance Evaluation

Selecting an appropriate classification model for the trash-sorting robotic arm was critical due to two key factors: the primary objective of accurate trash classification and the necessity of real-time operation. As more trash accumulates, fast detection and localization become increasingly vital to maintain efficient performance. Table 2 presents the performance metrics for three models: YOLOv8m, YOLOv8x, and YOLOv9m. From Table 2, it is evident that YOLOv8x demonstrates superior performance metrics compared to YOLOv8m and marginally outperforms YOLOv9m, which can be attributed to its higher complexity and deeper architecture. However, this comes with the drawback of significantly longer training times and the highest inference time for a standard 512x512 image. YOLOv9m outperforms YOLOv8m, despite their similar size, due to architectural improvements in YOLOv9. While YOLOv8 seeks accuracy and speed utilizing enhanced Path Aggregation Network (PANet) and Cross Stage Partial Network (CSPNet), YOLOv9 offers the Generalized Efficient Layer Aggregation Network (GELAN) and the Programmable Gradient Information (PGI), which reduces loss of information across layers, improving gradient stability and prediction accuracy [18]. Given the priority on real-time performance, YOLOv9m was selected for the prototype due to its balance between accuracy and fast detection.

Table 1. Performance comparison of YOLO models

Model	Inference time (ms)	Precision	Recall	mAP50	mAP50-95
YOLOv8m	18	0.764	0.773	0.706	0.476
YOLOv8x	25	0.825	0.814	0.779	0.553
YOLOv9m	17	0.791	0.803	0.755	0.511

This study focuses on accurately classifying trash within a pile, simulating realistic conditions a sorting robot might encounter, rather than using individual trash images with plain backgrounds. Figure 3 demonstrates the performance of the YOLOv9 model, which successfully detects various types of trash with high accuracy, even in complex, diverse backgrounds. The model reliably identifies challenging objects, such as bottles and attached caps. The detected objects, along with their locations and classes, were queued for processing. Upon receiving this data, the robotic arm initiated motion planning, efficiently calculating paths while accounting for joint constraints and obstacles (Fig. 4b), using RRT-Connect. The robot generated collision-free trajectories which were visualized in RViz (ROS 3d visualization program). It also calculated joint angles by IK to maintain the

Fig. 3: Classification and bounding box localization or trash type objects from test dataset using YOLOv9

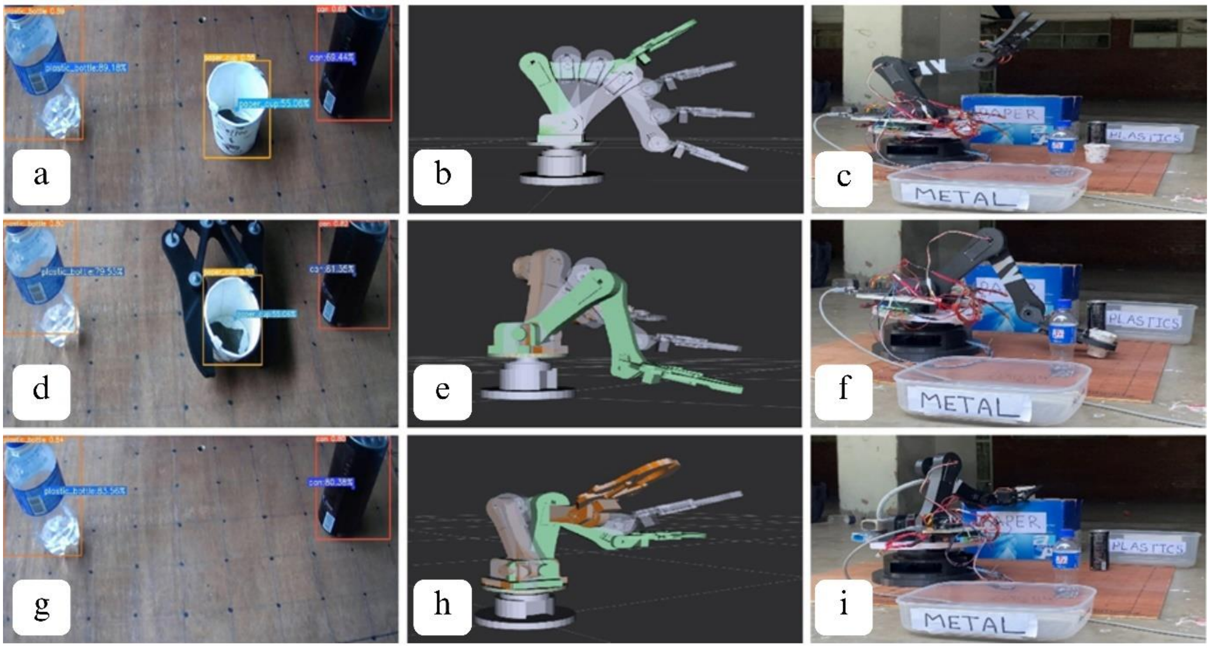

Fig 4. Implementation of trash sorting using a 4-DOF robotic arm. (a), (d), (g) Camera view of the trash with real-time detection. (b), (e), (h) Visualization in RViz (ROS 3d visualization program) during motion planning showing the robotic arm's current position (green), plan for the next pose (grey), and desired pose (orange). (c), (f), (i) Real-world performance of the robotic arm in trash sorting.

desired velocity and trajectory via a feedback loop. Real-life side views of different stages are depicted in Fig. 4(c, f, and i). After gripping an object (Fig. 4d), the robot planned and executed a path to deliver it to the appropriate bin (e.g., blue cardboard box in Fig. 4e). Upon successfully dropping each object, the robot returned to its default position and repeated the process until the task queue was empty (Fig. 4h).

Table 2 presents performance metrics for the robotic arm's trash sorting tasks in real-world conditions. The trash detection model achieved 83.5% accuracy with small mixed trash batches, though 24 misclassifications occurred, primarily between plastic and paper cups due to visual similarities. In 9 cases, paper items were undetected, often due to obstructions or color similarity with the platform. The robotic arm reached target locations with 85.86% accuracy, thanks to precise servo motor and sensor feedback, though slight deviations in the end-effector position were noted. A key challenge was grasping consistency, as the arm assumed all objects were at ground level, causing difficulty with variable heights. The two-clawed gripper handled diverse shapes but struggled with flat items and those larger than its open span. Despite these challenges, it achieved a 53.65% success rate in grasping, showing adaptability across materials. Once an item was gripped, the arm exhibited 92.05% accuracy in depositing trash into the correct bin, with occasional mis-drops mostly linked to sticky or irregularly shaped items. This high success rate was supported by the fixed bin positions, minimizing the need for dynamic adjustments during drop-offs. In summary, the experiment validates the algorithm's efficacy in integrating image detection, motion planning, and real-time trajectory generation (Fig. 4b, e, and h) for dynamic trash sorting tasks.

Table 2. Performance Analysis of the prototype 4-DoF trash sorting robotic arm

Metrics	Total Attempts	Successful Attempts	Success Rate
Trash Classification	200	167	83.5%
Reaching target coordinate after trash detection	191	164	85.86%
Successful Pickup after reaching target coordinate	164	88	53.65%
Dropping at the designated bin after picking up trash.	88	79	92.05%

4. Conclusion

This study implemented two highly developed algorithms for waste detection and motion planning of a 4-DoF robotic arm, specifically YOLOv9 and MoveIt, and evaluated the system's performance in real-world settings. Key findings include:

1. YOLOv9 demonstrated strong compatibility with precision, recall, and mAP metrics, achieving nearly 80% precision and recall, even with clumped or multiple trash objects. Moreover, Motion Planning via MoveIt achieved over 90% accuracy in reaching target coordinates, aided by servo motors and a closed-loop feedback system.
2. Limitations arose in object grasping, with the end-effector failing to pick up trash in nearly 50% of attempts due to the lack of height data and the limited grasping capability of the two-finger gripper. However, once grasped, the arm successfully deposited objects in the correct bins over 90% of the time.

This study reveals several avenues for future research, including trash object height detection for more effective trash grasping and integrating this with improved motion planning. Additionally, exploring the use of 5- or 6-DoF robotic arms could enhance dexterity, while developing more advanced end-effectors would enable better handling of a wider variety of trash objects. The insights from this study provide a foundation for advancing robotic waste management systems.

5. Acknowledgement

This study was conducted as a part of the course, Electro-Mechanical System Design and Practices (ME 361) at Bangladesh University of Engineering and Technology. We express our gratitude to our course instructors, Dr. Kazi Arafat Rahman and Priom Das for their guidance and supervision throughout this project.

6. References

[1] A. M. Madsen *et al.*, "Review of biological risks associated with the collection of municipal wastes," *Sci. Total Environ.*, vol. 791, p. 148287, Oct 2021.

[2] J. Petereit *et al.*, "ROBDEKON: Robotic Systems for Decontamination in Hazardous Environments," in *2019 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pp. 249–255, Sep. 2019.

[3] A. G. Satav, S. Kubade, C. Amrutkar, G. Arya, and A. H. Pawar, "A state-of-the-art review on robotics in waste sorting: scope and challenges," *International Journal on Interactive Design and Manufacturing*, vol. 17, issue 6, pp. 1-18, May 2023.

[4] X. Chen, H. Huang, Y. Liu, J. Li, and M. Liu, "Robot for automatic waste sorting on construction sites," *Autom. Constr.*, vol. 141, p. 104387, Sep. 2022.

[5] C. Liu *et al.*, "A Domestic Trash Detection Model Based on Improved YOLOX," *Sensors*, vol. 22, issue 18, pp. 6974, 2022.

[6] D. Patel, F. Patel, S. Patel, N. Patel, D. Shah, and V. Patel, "Garbage Detection using Advanced Object Detection Techniques," *2021 International Conference on Artificial Intelligence and Smart Systems*.

[7] I. Agustian, N. Daratha, R. Faurina, A. Suandi, and Sulistyaningsih, "Robot Manipulator Control with Inverse Kinematics PD-Pseudoinverse Jacobian and Forward Kinematics Denavit Hartenberg," *Jurnal Elektronika dan Telekomunikasi (JET)*, vol. 21, issue 1, Aug. 31, 2021.

[8] A. Al Tahtawi, M. Agni, and T. Hendrawati, "Small-scale Robot Arm Design with Pick and Place Mission Based on Inverse Kinematics," *J. Robot. Control JRC*, vol. 2, Nov. 2021.

[9] A. E. Ibrahim, R. Shoitan, M. M. Moussa, H. A. Elnemr, Y. I. Cho, and M. S. Abdallah, "Object Detection-based Automatic Waste Segregation using Robotic Arm," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, 2023.

[10] Y. Huang, "Research on garbage sorting robotic arm based on image vision," *J. Phys.: Conf. Ser.*, vol. 274, 2024.

[11] X. Liu, H. Wang, Y. Guo, and S. Qu, "Garbage sorting simulation of Panda robot arm based on deep learning and visual recognition", *Available: https://me336.ancorasir.com/wp-content/uploads/2023/06/Team2-Report.pdf*

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, June 2016.

[13] "Trash-detection image dataset." *Available: https://www.kaggle.com/datasets/ahnaftahmeed/trash-detection-image-dataset*

[14] Q. Liu, X. Wang, F. Liu, and L. Mu, "Research and Implementation of Robotic Arm Positioning and Grasping Based on Visual Guidance," *J. Phys.: Conf. Ser.*, vol. 2785, 2024.

[15] X. Gao *et al.*, "A rapidly exploring random tree optimization algorithm for space robotic manipulators guided by obstacle avoidance independent potential field," *International Journal of Advanced Robotic Systems*, June 26, 2018.

[16] K. Zou, X. Guan, Z. Li, H. Li, C. Jiang, and Z. Zhu, "Obstacle Avoidance Path Planning of a 4-DOF Weapon Arm Based on Improved RRT (RRT-H) Algorithm," *Mathematical Problems in Engineering*, vol. 2024, Issue 1, 2024.

[17] J. J. Kuffner and S. M. LaValle, "RRT-connect: An efficient approach to single-query path planning," *Proceedings of the 2000 IEEE International Conference on Robotics and Automation, ICRA 2000*, April 24-28, 2000.

[18] C. Y. Wang and H. Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," *arXiv preprint arXiv:2402.13616*, 2024.