VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2024, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 12, Number 2, April-June 2024 pp:164-173

Enhancing Efficiency and Safety with YOLOv5-Powered Robotic Arms for Waste Classification

Shaheena Noor [1*], **Muhammad Imran Saleem** [2], **Aneeta Siddiqui** [1], **Najma Ismat** [1], **Humera Noor Minhas** [3]

[1]Department of Computer Engineering, Sir Syed University of Engineering & Technology, Karachi - Pakistan; [2]Department of Software Engineering, Sir Syed University of Engineering & Technology, Karachi - Pakistan; [3]Quality Filtering Engineering Department, eyeo GmbH, Lichtstr., Cologne, 50825, North Rhine-Westphalia, Germany

Keywords: *Deep learning, YOLOv5, Robotic Arm, Object Picking*

Journal Info:
Submitted:
May 25, 2024
Accepted:
June 20 , 2024
Published:
June 30, 2024

Abstract The world is experiencing a transformation shift from manual labor to digital solutions, making work simpler and more efficient while enhancing the quality of life globally. A prime example of this shift is the Object Picking Robotic Arm (OPRA). Designed to operate with minimal human intervention, the OPRA reduces the risk of physical injuries among workers by replacing human labor with robotic precision. This technology finds applications in both industrial and domestic settings, including the automotive industry, metalworking, chemical processing, and various pick-and-place tasks. In this research, we develop a robotic system for automated waste picking and sorting. This system utilizes the YOLOv5 object detection algorithm to achieve high accuracy (95%) and precision (90%) in classifying five common waste categories: cardboard, metal, paper, plastic, and trash.

***Correspondence author email address:** shanoor@ssuet.edu.pk
DOI: 10.21015/vtse.v12i2.1813

1 Introduction

Robotic arms are indispensable in industries where efficiency is critical, particularly in environments hazardous to human health or those requiring the simplification of complex tasks. Due to the necessity of actuators and precise component movements, most industrial robotic arms are expensive. However, by advancing research to reduce costs and enhance their utility, we can increase the flexibility of robotic arms, thereby promoting the automation of manual labor [1]. However, reducing costs often necessitates the use of mid-grade components and making certain compromises in the market [2].

The pricing of robotic arms is influenced by several factors, including speed, cost, repeatability, human safety, and workload capacity. Crucial aspects include the arm's ability to grip items effectively and its high repeatability, ensuring it consistently returns to its home position after completing tasks [3] [4], [5].

The demand for interior service robots is growing due to two significant social challenges: an aging population and rising labor costs. To perform various grasping tasks, service robots in homes or workplaces must identify target items within complex or changing environments. Developing a lightweight recognition algorithm capable of distinguishing different objects under varying conditions—such as illumination, occlusion, and object posture—while ensuring real-time responsiveness and accurate gripping is highly challenging.

Recent robotics research has led to the development of numerous alternative gripping techniques. Deep learning algorithms, particularly those utilizing Convolutional Neural Networks (CNNs), have become the most prominent in grasp synthesis. These techniques require extensive data, substantial training time, and costly hardware environments. Despite their potential, these methods often suffer from overfitting, limited generalizability, and complexity in interpretation. The diverse target items, varying viewing angles, and dynamic environments typical of indoor robotic grasp tasks further complicate the implementation of deep learning-based approaches [6].

In today's world, humans and robots collaborate to complete tasks. Autonomous, self-managing robots—referred to as aid robots—are now employed across various sectors, including business, military, and healthcare. For instance, robots can replace humans in hazardous tasks such as working in high-temperature environments or defusing explosives [6].

The successful use and implementation of robotic arm sorting technology in American and European industries have driven this project. These regions have effectively leveraged robotics to help labor forces meet deadlines impacted by manual tasks. However, this technology has not yet been widely adopted in Asian countries. A particularly successful initiative is currently operating across America and Europe, significantly aiding the workforce and impacting society by enabling job automation according to industry capabilities [6].

Humanoid grasping is vital for developing humanoid robots with anthropomorphic hands. It presents [7] a deep learning-based control framework that uses dynamic contact data to enhance grasping efficiency. By utilizing an under actuated hand based on human hand data and employing hand gestures to reduce control complexity, the system outperforms static analysis methods. Tested in both simulations and real robots, the framework achieves over 90% success in grasping various objects, including thin items like cards. This approach effectively addresses high DoF control challenges and expands grasp capabilities, offering a promising direction for future research on anthropomorphic hands.

The paper is arranged as follows: Section II covers the detailed literature study on object detection in different domain, object tracking and segmentation. Sections III discusses the theoretical and mathematical foundations of YOLOv5 architecture. We give an overview to the experimental platform in Section IV, experiments and results in Section V and conclude the paper in Section VI.

2 Literature Review

This section reviews current advancements in the field of trash-picking robots. The tasks of trash collection and disposal are tedious and repetitive, highlighting the need for robotic solutions to improve efficiency and reduce human labor [8]. Every year, over 14 billion pounds of garbage are dumped into the world's oceans, causing significant environmental damage and highlighting the urgent need for effective waste management solutions [9]. Workers tasked with trash collection and disposal face health risks, including heart disease and metabolic syndrome, underscoring the importance of implementing safer and more efficient waste management practices [10]. Implementing automatic systems, primarily involving robots and processing units like Arduino and Raspberry Pi, can help prevent these health risks by reducing the need for manual labor in waste collection and disposal [11].

Robots lacking the ability to identify trash can be harmful. Therefore, an intelligent garbage bin robot, utilizing a robot operating system for waste collection, is essential for efficient and safe waste management [12]. A smart garbage classification system [13, 14] is highly beneficial, not only for worker safety but also for economic efficiency in recycling. Typically, discarded waste consists of 5.8% metals, 3.5% glass, 1.6% plastic, 12.9% paper, 1.8% textiles, and 53.7% biodegradable materials [15, 16]. Extensive work has been done on garbage detection using image processing and deep learning. Image processing techniques are employed for feature extraction, followed by the use of classifiers to identify various objects [17]. The study of IoT-based automatic garbage collection using smart robots is explored in [18]. Similarly, an IoT-based smart waste bin monitoring and management system is proposed in [19]. In [20], an autonomous robot designed to pick up trash from grass using a deep neural network is proposed. Utilizing the GINI dataset with the GarbNet network for garbage detection, it achieved a mean accuracy of 87.69% [21]. In [22], a comparative study evaluates the performance of Mask Region-Based Convolutional Neural Network (Mask-RCNN), You Only Look Once (YOLOv4), and YOLOv4-tiny for a trash-collecting robot using the Rocker-Bogie mechanism. The experimental results show that Mask-RCNN achieved a Mean Average Precision (mAP) of 83%, YOLOv4 achieved 97.1% mAP, and YOLOv4-tiny achieved 95.2% mAP. These results were specifically for detecting plastic materials.

In [23], a deep learning-based robot was designed for garbage collection, specifically targeting cans, plastic bottles, and lunch boxes, achieving 95.6% precision and 96.8% recall.

The underwater garbage cleaning robot designed in [24] utilizes YOLOv4 with a Mean Average Precision (mAP) of 95.099%. Additionally, in [25], the YOLOv3 network is employed to collect floating plastic garbage.

From the above literature review, we have some research questions as given below:

- **Cost vs. Functionality Trade-off in Robotic Arms:** How can we design cost-effective robotic arms for industrial applications while maintaining essential functionalities like gripping and repeatability?
- **Lightweight Object Recognition for Service Robots:** Can we develop lightweight object recognition algorithms using deep learning that are suitable for service robots operating in real-time with limited computational power and under varying environmental conditions?
- **Overcoming Limitations of Deep Learning for Grasping Tasks:** How can we address challenges like overfitting, limited generalizability, and complexity in deep learning-based grasping techniques, particularly for diverse objects and dynamic indoor environments faced by service robots?
- **Adoption of Robotic Arm Sorting Technology in Asia:** What are the factors hindering the wider adoption of robotic arm sorting technology in Asian countries compared to America and Europe?

This research paper focuses on developing lightweight object recognition algorithm using deep learning, tailored for service robots operating in real-time with limited computational resources and diverse environmental conditions. The goal is to mitigate the challenges posed by constrained computational power while ensuring robust performance in dynamic environments with varying lighting and other factors, empowering service robots to operate effectively.

2.1 Objective of Research

This research aims to develop a robotic system for automated waste picking and sorting. It utilizes a deep learning model based on the YOLOv5 algorithm in classifying five common waste categories (cardboard, metal, paper, plastic, and trash). This technology has the potential to improve safety and efficiency in waste management by reducing manual labor and enabling robots to handle waste sorting tasks.

2.2 Novelty of Work

This research presents a novel application of the well-known YOLOv5 object detection algorithm. By utilizing YOLOv5 for robotic waste picking, the study integrates computer vision and robotics for a practical task in waste management. Additionally, the research focuses on classifying five specific waste categories, which could be novel if existing work deals with broader waste object categories or different environments. Furthermore, achieving high accuracy and precision (95% and 90% respectively) for these specific waste types through YOLOv5 optimization could be a significant contribution compared to previous attempts in this area.

3 Methodology

According to Edmundo Casas [26], although YOLOv8 is the latest variant, YOLOv5 emerges as the better choice due to several key factors. First is its **Superior Efficiency**, YOLOv5 shines in efficiency metrics. It achieves higher precision, recall, and F1-scores while requiring less training time and fewer epochs to reach optimal performance. This efficiency makes YOLOv5 ideal for scenarios with limited resources or tight deadlines. Second is its **Faster Inference**, for real-time applications, YOLOv5 boasts faster inference times, allowing for quicker object detection. While YOLOv8's inference speeds are respectable, YOLOv5's edge proves advantageous in time-sensitive situations. Therefore, considering these advantages and the specific needs of our project, we opted to utilize YOLOv5 for our object detection tasks.

3.1 YOLOv5 Network Architecture

In the YOLO architecture series, the latest iteration is YOLOv5 [27], which boasts the fastest detection speed of up to 140 frames per second. This network offers very fast inference speeds compared to other architectures, while maintaining high detection accuracy. Notably, the total size of the weight file for the YOLOv5 model is about 90% smaller than that of YOLOv4, making it highly suitable for real-time object detection applications.

The object detection algorithm based on deep learning has been extensively studied in various domains such as agriculture [28], medical diagnosis [29], transportation [30], among others. However, there is relatively limited research in the domain of robotics. Therefore, this paper proposes the use of the YOLOv5 network to enhance the flexibility of object recognition in robots and improve localization accuracy. The key points of the YOLOv5 is given below:

1. Backbone Network: YOLOv5 uses the CSPDarknet53 backbone network, which is a variant of Darknet. CSPDarknet53 includes Cross Stage Partial connections to improve information flow and gradient propagation.
2. Neck: YOLOv5 utilizes a feature aggregation module called PANet (Path Aggregation Network). PANet helps in combining features from different layers with different resolutions, thereby improving the detection performance.
3. Detection Head: The detection head of YOLOv5 consists of three prediction heads (YOLOv5s) or five prediction heads (YOLOv5x). These prediction heads predict bounding boxes, objectness scores, and class probabilities at three different scales.
4. Training: YOLOv5 is trained using the COCO dataset and optimized using the Adam optimizer. It employs data augmentation techniques such as mosaic data augmentation, random scaling, translation, flipping, color distortion, and Gaussian augmentation.
5. Inference: During inference, YOLOv5 applies non-maximum suppression (NMS) to filter redundant detections based on the confidence scores of the predicted bounding boxes.
6. Speed and Performance: YOLOv5 achieves real-time object detection with high accuracy. It offers various model sizes (YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x) that trade off between speed and accuracy, with YOLOv5s being the smallest and fastest, and YOLOv5x being the largest and most accurate.
7. Model Size: YOLOv5 models have relatively smaller model sizes compared to YOLOv4, making them suitable for deployment on devices with limited computational resources.
8. Community and Development: YOLOv5 is an open-source model with a vibrant community for support and development. The model is accessible on GitHub, allowing for customization and fine-tuning tailored to specific applications.

This summary provides an overview of the architecture of YOLOv5, highlighting its essential components and features. YOLOv5 has gained popularity due to its effective balance between speed and accuracy in real-time object detection tasks.

3.2 YOLOv5 Object Detection

This section provides a detailed exploration of the mathematical foundation underlying the YOLOv5 architecture for object detection. It offers a structured overview of the algorithmic workings of YOLOv5 in object detection tasks.

Algorithm 1 outlines the main steps involved in the YOLOv5 object detection process, from input image processing through to the calculation of losses and final detections using non-maximum suppression.

4 Experimental Platform

The machine vision-based robot object detection system comprises three layers, as illustrated in Figure 1:

1. Input Layer
2. Control Layer
3. Output Layer

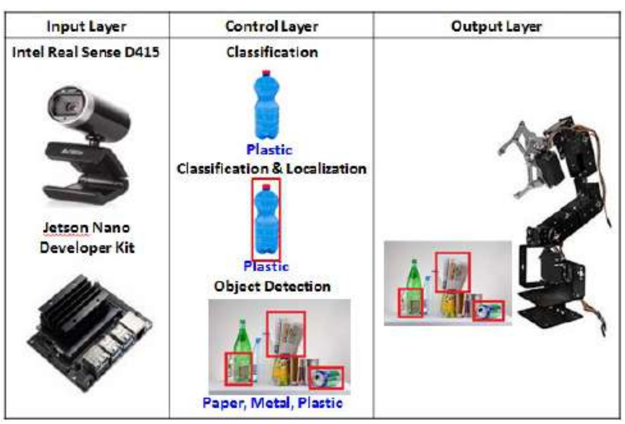

Figure 1. Elements of Object Grasping Robotic Arm

All server control terminal and foreground control terminal operations are managed through the input layer. The control layer handles all critical systems and controls, including an image processing system, image acquisition system, object recognition system, and robot control system. The output layer manages the execution and image acquisition equipment. This research introduces a robot object detection system

Algorithm 1. YOLOv5 Object Detection Pipeline

Input:

Image: Image input for processing.

Backbone Network (CSPDarknet53):

Input: Image

Output: Feature maps

feature_maps = CSPDarknet53(image)

Feature Pyramid Network (FPN):

Input: Feature maps from CSPDarknet53

Output: FPN features

fpn_features = FPN(feature_maps)

Object Detection Head (YOLO Head):

Input: FPN features

Output: Predicted bounding boxes, objectness scores, and class probabilities.

predictions = YOLO_Head(fpn_features)

Bounding Box Regression:

$$(b_x, b_y, b_w, b_h) = sigmoid(predictions[:, :4]) + Anchor$$

where Anchor represents the Predefined anchor boxes

Objectness Score:

objectness_scores = sigmoid(predictions[:, 4])

Class Prediction:

class_scores = softmax(predictions[:, 5:])

where sigmoid represents the Sigmoid activation function softmax represents the Softmax activation function

Loss Calculation: Bounding Box Regression Loss:

$$loss\_box = \lambda\_box \cdot MSE(predictions[:, :4], targets[:, :4])$$

Objectness Loss:

$$loss\_obj = \lambda\_obj \cdot BCE(sigmoid(predictions[:, 4]), targets[:, 4])$$

Classification Loss:

$$loss\_cls = \lambda\_cls \cdot CE(softmax(predictions[:, 5:]), targets[:, 5:])$$

Total Loss:

$$total\_loss = loss\_box + loss\_obj + loss\_cls$$

where
MSE is the Mean Squared Error
BCE is the Binary Cross-Entropy
CE is the Cross-Entropy
$\lambda\_box, \lambda\_obj, \lambda\_cls$: are the Loss weights

Non-Maximum Suppression (NMS):

Input: Predicted bounding boxes, objectness scores, and class probabilities

Output: Final detection following NMS

detection = NMS(predictions)

designed to incorporate a robotic arm for grasping and object localization, as depicted in Figure 2. The robotic arm setup consists of an Xarm, a detection platform, an Intel RealSense D415 camera, and a server. The Xarm is an affordable and lightweight programmable device that includes a robotic arm, a control cabinet, signal cables, a power supply, and other essential components. The Intel RealSense D415 camera features a D410 depth sensor, providing high-resolution and precise depth sensing capabilities. It converts optical and electrical signals and transmits them to the server for further processing.

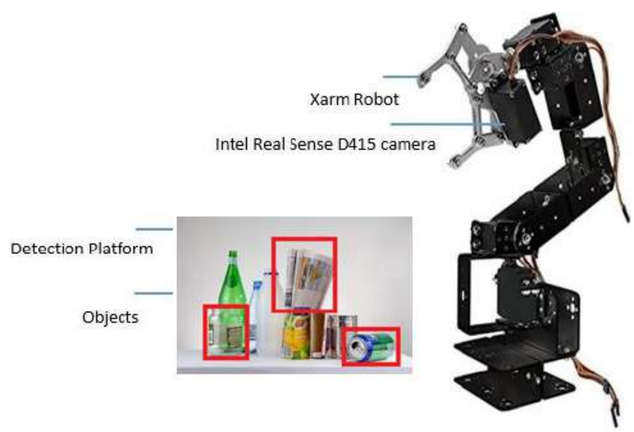

Figure 2. Robotic Vision System for Object Detection

This configuration empowers the robotic arm to execute precise grasping and localization tasks, harnessing the capabilities of the Xarm and Intel RealSense D415 camera for accurate object detection and manipulation.

4.1 Experimental Process

Object detection, is the central theme of this research, that encompassing both object localization and classification. Object classification identifies the correct category of an object and determines its spatial location. The robotic arm with grasping capabilities utilizes three key functionalities: object grasping, classification, and localization. Figure 3 depicts the experimental setup, enabling the robotic arm to effectively detect, classify, and localize objects. The figure details how the robotic arm performs these tasks.

Object Classification: To begin, we collected and

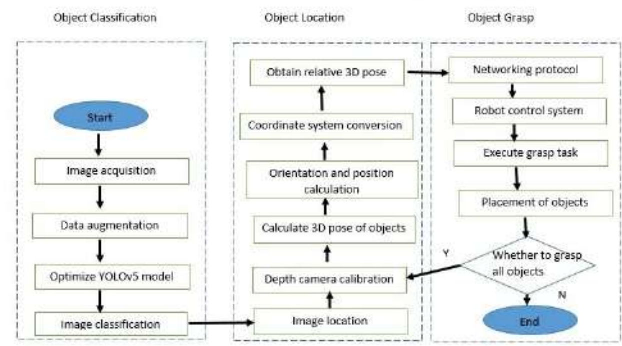

Figure 3. OPRA: Object Manipulation Process

augmented object images to expand our collection. These augmented images, along with the originals, were included in our dataset. Following that, we trained the YOLOv5 network on this dataset. We then categorized the object images and performed additional improvements to improve the precision and accuracy of our item categorization.

This approach involved acquiring and enhancing photos, incorporating them into the dataset, training the YOLOv5 network for accurate object recognition, and then fine-tuning the classification accuracy and precision.

Object Localization: We captured images of objects and used a depth camera to gather depth information. Next, we calculated the 3D pose of each object. To align different coordinate systems, we employed the Hand-Eye calibration method, which established a pose conversion matrix. This matrix was then used to accurately determine the object's three-dimensional position relative to the robot's location.

This approach ensures precise spatial localization of objects within the robot's surroundings, allowing for effective interaction and manipulation tasks.

Object Grasping Initially, we established a secure communication link between the robot and the host using the TCP/IP protocol [31]. Following this, the robot control system received the three-dimensional poses of each object image. Guided by the robot control system, the robot proceeded to grasp the specified object. Once grasped, the robot placed the object in its designated position.

This process encompassed setting up a secure communication link, receiving object poses, guiding the robot to grasp the object, and accurately placing it in its intended location. It ensures efficient and precise handling of objects by the robot.

5 Experiments and Results

5.1 Dataset

The dataset used in this study was sourced from Kaggle (https://www.kaggle.com/datasets/techsash/waste-classification-data) to evaluate our results. It consists of five classes: cardboard, metal, paper, plastic, and trash. While Kaggle datasets offer accessibility and benchmarking advantages, they may not fully capture local waste characteristics. To address this, the study analyzed local household waste, segregated it into relevant classes, and automated the process using IoT and machine learning. Object recognition capabilities of the YOLOv5 model were leveraged to accurately identify and classify waste items in real-time. Details of the dataset are provided in Table 1. We split the dataset into 85% for training and 15% for testing. A subset of the data is illustrated in Figure 4.

Table 1. Dataset Details

Category	Number of Images	Train	Test
Cardboard	403	282	121
Metal	410	287	123
Paper	594	415	178
Plastic	482	337	145
Trash	137	96	41

Figure 4. Subset Details of Waste Classification Dataset

5.2 Results

For waste classification, we achieved a precision of 90.89% and an accuracy of 95.61% using YOLOv5. Table 2 provides detailed measures for each category, demonstrating the effectiveness of our approach in accurately identifying and classifying waste materials. These results highlight the robustness of YOLOv5 in handling diverse waste types, including cardboard, metal, paper, plastic, and trash. The high precision and accuracy underscore the potential of our method for practical applications in waste management and environmental monitoring. Consider Figure 5, that depicts the confusion matrix used to forecast various waste classification.

Table 2. Waste Classification Results

Category	Precision(%)	Accuracy(%)
Cardboard	88.23	95.23
Metal	94.11	97.6
Paper	96.15	98
Plastic	93	95.12
Trash	77	92.10
Total	90.89	95.61

	Cardboard	Metal	Paper	Plastic	Trash
Cardboard	95	0	3	0	2
Metal	0	98	0	2	0
Paper	0	0	98	0	2
Plastic	1	1	3	95	0
Trash	2	0	3	3	92

Figure 5. Confusion Matrix of Waste Classification

5.3 Comparative Analysis

In comparison [22] to the YOLOv4-based system using Raspberry Pi 4 B+ for trash bin automation, which achieved 75% accuracy with YOLOv4-tiny and 91% with full YOLOv4 on a PC, our research with YOLOv5 has achieved higher accuracy and precision. Specifically, the YOLOv5 model classifies five categories of daily waste: cardboard, metal, paper, plastic, and trash, achieving 95% accuracy and 90% precision. This highlights the superior performance and efficiency of YOLOv5 in categorizing a broader range of waste types compared to the YOLOv4-based system.

6 Discussion and conclusions

Industrial robotic arms are pivotal in various sectors due to their ability to perform tasks ranging from simple to complex, across diverse environmental conditions. Automating these tasks not only enhances safety by reducing human exposure to hazardous environments but also frees up human resources for more valuable and strategic activities, such as customer engagement.

In this study, we have implemented a deep learning-based garbage classification model using YOLOv5. This model focuses on categorizing five common types of daily waste: cardboard, metal, paper, plastic, and trash. Utilizing standard datasets, YOLOv5 was constructed and trained, resulting in experimental outcomes showing a commendable 95% accuracy and 90% precision.

Looking ahead, future research directions could explore the use of larger and more diverse datasets. With increased data volume, the model's performance is expected to improve, enabling it to detect and classify an even wider range of waste classes in real-time scenarios. This approach holds promise for enhancing waste management practices and environmental monitoring efforts.

Author Contributions

Shaheena Noor: Conceptualization and designed the study, developed the methodology, conducted the experiments, and analyzed the data **Muhammad Imran Saleem**: assisted in refining the methodology, provided critical feedback on experimental design, and contributed to data analysis. **Aneeta Siddiqui**: Visualization, Exploring of data **Najma Ismat**: Software, Validation. **Humera Noor Minhas**: contributed to the interpretation of results and the writing of the manuscript.

Compliance with Ethical Standards

The authors declare no conflict of interest. This article does not involve studies with human participants or animals conducted by any of the authors. Informed consent was obtained from all individual participants included in the study.

References

[1] K. E. Clothier and Y. Shang. A geometric approach for robotic arm kinematics with hardware design, electrical design, and implementation. *Journal of Robotics*, 2010, 2010.

[2] M. A. Rahman, A. H. Khan, T. Ahmed, and M. Sajjad. Design, analysis and implementation of a robotic arm-the animator. *American Journal of Engineering Research (AJER)*, 2(10):298–307, 2013.

[3] A. Zade R. Gautam, A. Gedam and A. Mahawadiwar. Review on development of industrial robotic arm. *International Research Journal of Engineering and Technology (IRJET)*, 4, 2017.

[4] Last seen on 5 July 2024. How much does a robotic arm cost?, 2024.

[5] Last seen on: 07 July 2024. Recent developments in robotics, 2022.

[6] Y. Jadeja and B. Pandya. Design and development of 5-dof robotic arm manipulators. *International Journal Of Scientific & Technology Research*, 8(11):2158–2167, 2019.

[7] Shaowen Cheng, Yongbin Jin, and Hongtao Wang. Deep learning-based control framework for dynamic contact processes in humanoid grasping. *Frontiers in Neurorobotics*, 18, 2024.

[8] J. Palacin, J. A. Salse, I. Valgano ́n, and X. Clua. Building a mobile robot for a floor-cleaning operation in domestic environments. *IEEE Transactions on Instrumentation and Measurement*, 53(5):1418–1424, 2004.

[9] G. Sharma, B. Sinha, Pallavi, H. Hakkim, B. P. Chandra, A. Kumar, and V. Sinha. Gridded emissions of CO, NOx, SO2, CO2, NH3, HCL, CH4, PM2.5, PM10, BC, and NMVOC from open municipal waste burning in india. *Environmental Science & Technology*, 53(9):4765–4774, 2019.

[10] Y. Demiral, A. Soysal, A. C. Bilgin, B. KiLic ̧, B. Unal, R. UcKu, and T. Theorell. The association of job strain with coronary heart disease and metabolic syndrome in municipal workers in turkey. *Journal of occupational health*, 48(5):332–338, 2006.

[11] S. S. Chandra, M. Kulshreshtha, and P. Randhawa. A review of trash collecting and cleaning robots. In *9th International Conference on Reliability, Infocom Technologies*

and Optimization (Trends and Future Directions)(ICRITO), pages 1–5. IEEE, 2021.

[12] W. S. Cheong, S. F. Kamarulzaman, and M. A. Rahman. Implementation of robot operating system in smart garbage bin robot with obstacle avoidance system. In *Emerging Technology in Computing, Communication and Electronics (ETCCE)*, pages 1–6. IEEE, 2020.

[13] S. Hossain, B. Debnath, A. Anika, M. Junaed-Al-Hossain, S. Biswas, and C. Shahnaz. Autonomous trash collector based on object detection using deep neural network. In *2019 IEEE Region 10 Conference (TENCON)*, TENCON, pages 1406–1410. IEEE, 2019.

[14] B. Adarsh, H. Shivayogappa, and K. et al. Navya. Automated smart sericulture system based on 6lowpan and image processing technique. In *International Conference on Computer Communication and Informatics (ICCCI)*, ICCCI, pages 1–6. IEEE, 2016.

[15] R. Rokade, A. Maurya, V. Khade, and J. et al. Mali. Smart garbage separation robot with image processing technique. *International Journal of Engineering Research & Technology*, (Special Issue), 2018.

[16] G. S. Kannan, S. S. Kumar, R. Ragavan, and M. Balakrishnan. Automatic garbage separation robot using image processing technique. *International Journal of Scientific and Research Publications*, 6(4):326–328, 2016.

[17] J. S. Denker D. Henderson R. E. Howard W. Hubbard Y. LeCun, B. Boser and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1:541551, 1989.

[18] M. Sathveek, C. Shreemanth, J. S. Asish, H. T. Agalayil, I. T. B. Mary, and J. J. Paul. Iot based automatic garbage collection for smart cities. In *6th International Conference on Devices, Circuits and Systems (ICDCS)*, pages 477–479. IEEE, 2022.

[19] T. Ali, M. Irfan, A. S. Alwadie, and A. Glowacz. Iot-based smart waste bin monitoring and municipal solid waste management system for smart cities. *Arabian Journal for Science and Engineering*, 45(12):10185–10198, 2020.

[20] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu. Deep learning based robot for automatically picking up garbage on the grass. *IEEE Transactions on Consumer Electronics*, 64(3):382–389, 2018.

[21] M. Yang and G. Thung. Classification of trash for recyclability status. CS229 project report 2016, 2016.

[22] Aria Bisma Wahyutama and Mintae Hwang. Yolo-based object detection for separate collection of recyclables and capacity monitoring of trash bins. *Electronics*, 11(9): 1323, 2022.

[23] Y. Arai, R. Miyagusuku, and K. Ozaki. Development and testing of garbage detection for autonomous robots in outdoor environments. In *IEEE/SICE International Symposium on System Integration (SII)*, pages 138–143. IEEE, 2021.

[24] M. Tian, X. Li, S. Kong, L. Wu, and J. Yu. A modified yolov4 detection method for a vision-based underwater garbage cleaning robot. *Frontiers of Information Technology & Electronic Engineering*, 23(8):1217–1228, 2022.

[25] S. Kong, M. Tian, C. Qiu, Z. Wu, and J. Yu. Iwscr: An intelligent water surface cleaner robot for collecting floating garbage. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(10):6358–6368, 2020.

[26] Edmundo Casas, Leo Ramos, Eduardo Bendek, and Francklin Rivas-Echeverria. Yolov5 vs. yolov8: Performance benchmarking in wildfire and smoke detection scenarios. *Journal of Image and Graphics*, 12(2):127–136, 2024.

[27] Y. Liu, B. Lu, J. Peng, and Z. Zhang. Research on the use of yolov5 object detection algorithm in mask wearing recognition. *World Scientific Research Journal*, 6(11):276–284, 2020.

[28] S. Butte, A. Vakanski, K. Duellman, H. Wang, and A. Mirkouei. Potato crop stress identification in aerial images using deep learning-based object detection. *Agronomy Journal*, 113(5):3991–4002, 2021.

[29] D. W. Wong, A. P. Yow, B. Tan, Y. Xinwen, J. Chua, and L. Schmetterer. Localization of anatomical features in vascular-enhanced enface oct images. In *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 1875–1878. IEEE, 2020.

[30] M. Haris and A. Glowacz. Road object detection: A comparative study of deep learning-based algorithms. *Electronics*, 10(16):1932, 2021.

[31] A. Mejías, R. S. Herrera, M. A. Márquez, A. J. Calderón, I. González, and J. M. Andújar. Easy handling of sensors and actuators over tcp/ip networks by open source hardware/software. *Sensors*, 17(1):94, 2017.