GreenSort: A Vision-Centric Robotic System for Smart Waste Collection and Environmental Protection using SSD MobileNet V2 and the TACO Dataset

Andrea Amalanathan[1], *Saurav* Pampana[1], *Kshitij*[1], and *Anita Christaline* Johnvictor[2*]

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
[2]Centre for Neuro Informatics, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

Abstract. As part of sustainability, care for the environment is necessary to ensure health of everyone. Sadly, pollution and particularly stray waste have turned out to be a major problem for humans and wildlife. All the litter, which is not properly disposed, adds to the destruction of the environment. This research discusses a new solution to this problem: a trash picking robot. The proposed robot uses a visual collection system to detect and collect nearby trash, making it a great tool for multi-area cleaning. Robotics has changed our lives in many ways, saving time, money, and helping in many areas. This robot is designed to analyse the size, shape, and position of items to accurately, pick them up and transfer them to different types of waste classification. The implementation begins by identifying pollution problems and exploring existing solutions. Multiclass segregation of different types of wastes using SSD MobileNet V2 on TACO Dataset is performed. The TensorFlow model provides an accuracy of 96%. Ultimately, the goal is to show that technology plays an important role in keeping our world clean and green.

1 Introduction

Waste on earth seemingly came with humanity, but its effects have been enormous and has attracted serious attention in recent times. We are usually eager to pay to dispose garbage. In case there is no spending involved in disposal, then it would be a valuable commodity. Garbage was a serious problem in big cities in the middle ages but not as dense as it is nowadays. With the growing population density and the growing rates of individual pollution, the waste disposal issue may become even more important in the future.

---

*Corresponding author: anitachristaline.j@vit.ac.in

Figure 1 shows the trends of waste generated over the years, for instance, in 1960, solid waste production was about 88 million tons worldwide, which shot up to a high of 253 million tons in 2005 then slightly about 251 million tons in 2012. Unless we address pollution today, it might become a bigger problem to our upcoming generations.

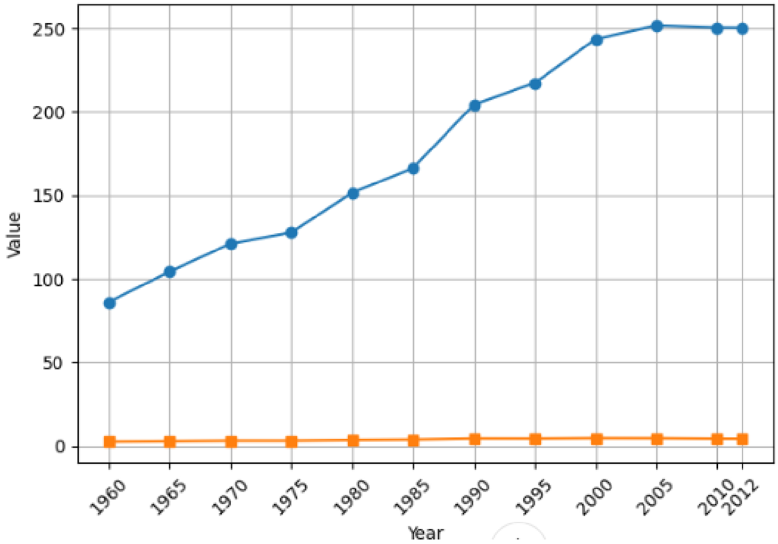


Fig. 1. Trends of waste generated over the years

Nevertheless, different wastes have different durations of decomposition as shown in Fig.2. For instance, paper has a degradation life of three months, whereas glass bottles have a degradation life of up to 4,000 years. The synthetic and chemical growth has been rapid. These waste also puts the earth in more strain to control pollution and underlines the need for human intervention. Waste management involves many individuals such as sanitation workers and community volunteers, but the quantity of waste is increasing faster day by day. This pollution is not only a threat to ecological health, but also to human psychological well-being.

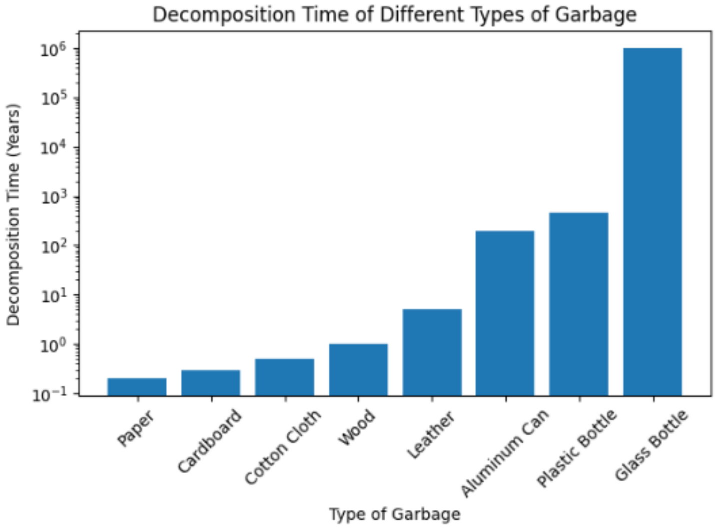


Fig. 2. Decomposition time for different types of wastes

Studies have shown that the level of stress and depression can rise among the residents due to the exposure to polluted environments. Robots hold a potential solution to this increasing menace. Robots are increasingly becoming multifaceted and are currently used in such industries as manufacturing, medicine, and the military. Robots are not exhausted, do not need payments and do not need breaks like human workers. Therefore, their use in collecting and disposing waste can be very efficient.

Potentially, robots could be utilized to clean up different environments. Equipped with sensors and intelligent software, such robots would be capable of identifying various wastes, collect litter and carry them to dumping sites. For instance, recognizing and processing an object like a plastic bottle, on the basis of its material and shape would enable robots to aid in the segregation and recycling of waste. These selfless workers would make the world a cleaner and healthier place in the future. In subsequent sections, we shall elaborate on how we have investigated the pollution problem, examined the available waste management systems and outline our robotics system for garbage collection.

2 Literature Survey

A study proposed by Murakami and Goto contributed a lot in providing a novel perspective of robotic grasping through involving deep learning and tactile feedback in autonomous systems. Their work stressed on the fact that sensory perception and machine learning combined algorithms make robot more adaptive for unstructured environment. The study presented a convolutional neural network (CNN)-based model that is trained using a variety of tactile. The researchers emphasized that visual reviews were not enough to aid in precise manipulation activities; therefore, tactile sensing gave a further context on; object texture, hardness and slippage. This study provided the groundwork in multi-modal perception of robots through the combination of deep learning and tactile feedback forming a vital stepping stone for later hybrid vision-tactile manipulation studies. [1].

An adaptive reinforcement learning framework aimed at improving robotic manipulation through self-learning and feedback optimization mechanisms [2]. Their system integrated deep reinforcement learning (DRL) with robotic control theory to enable continuous skill acquisition. Unlike static learning approaches, their model dynamically updated policy parameters based on environmental interactions, thereby enabling the robot to adjust its grasping force, trajectory, and angle autonomously. Nevertheless, the system consumed significant computational resources to run in real time and additional optimization of the reward shaping was recommended to enhance the convergence rate and stability in the large scale implementation.

An industrial-grade waste sorting system that used computer vision and robotic automation was presented and applied to the recycling pro- cesses of waste [3]. The paper presented a 6-axis robotic arm with high-definition cameras and neural network-based classifiers that have the ability to scan and sort plastic, glass, metal, and paper with a high accuracy rate to up to 99 percent. Convolutional neural networks that had been trained on several thousands of labelled images of waste items were used in the vision pipeline, which guaranteed a high level of detection even when the conditions were unfavourable in terms of lighting and contamination. Their system had real-time decision-making algorithms, which identified the type of material and sent robotic actuators to pick and place it in an efficient manner.

Research show how deep learning models can be applied in the classification of materials used in waste management robotics [4]. They used convolutional neural networks (CNNs) to extract features and classify mixed-material waste in their study, and high intra-class variability materials, namely transparent plastics and metallic foils, were used. Their

ResNet based model was optimized, and it scored more than 94 percent on the test data which is significantly higher than the traditional machine learning classifiers. Significant drawbacks however were the high computational requirement and low inference on embedded systems indicating the necessity of light-weight architecture like MobileNet to be used in real-time applications.

A vision-based grasping system was developed by Kragic and Bjorkman incorporated monocular and binocular images of various cameras to provide correct 3D perceptions and object manipulations in the robots [5]. Their method was one of the first to attempt to reconcile perception and control in robotics using the visual servoing techniques. Their experiments of the standard grasping benchmarks found that they improved substantially on their precision particularly when using small and partially obscured objects. Another concept that was mentioned by Kragic and Bjorkman is that of feedback loops, which allows the robot to dynamically correct pose error during execution of grasps. This work was the foundations of the future 3D vision-based manipulation systems and contributed to subsequent literature in the area of active perception and camera-in-hand control paradigms.

One of the key improvements of the study was the consideration of the obstacle avoidance, which meant that the planned grasps could be performed in the cluttered environment without colliding. The authors suggested that it would be a good idea to exploit parallel computing and heuristic optimization to enhance efficiency. This work had a potent impact on the research later in grasp simulation, introduction of haptic feedback, and grasp refinement during learning. The first appearance of such approach was given by Livine et al where they combined optimistic control theory with the concept of experiential learning allowing the robots to learn grasping skills with the help of the iterative feedback [6]. Their system integrated both the reinforcement cues provided by sensory input and predictive analysis of motor control commands which the robot could learn to optimize its motion control policies with time. Their model connected the gap between analytical control and data-driven learning and positively shaped the evolution of modern algorithms of reinforcement learning in robotics.

Kalashnikov et al. developed robotic grasping with a combination of deep learning and haptics viewpoint, creating one of the first multimodal learning frameworks on the task of a manipulator trade-off [7]. The authors suggested a neural network framework that trained to estimate the likelihood of grasping success using visual and sensory inputs of tactile feedback. The visual data was cut out through RGB images, whereas the data of the touch was recorded by means of the pressure and force sensors inside the gripper. Through the collective maximization of the two modalities, the model performed better than the vision-only systems especially with the presence of occlusions or ambiguous objects.

The high-precision robotic arm researched by Lahoti et al and was specifically meant to be used in the collection and sorting of wastes, which can be seen as a viable move to the automation of industrial scale in the field of environmental robotics [8]. The experiment was aimed at the optimization of mechanical design and precision of control, where a computer-vision feedback was incorporated to detect the waste correctly and complete the pick and place operations. The robotic arm had multi-degree-of-freedom actuation, energy-efficient lightweight materials, and torque sensors to ensure a consistent gripping force according to the type of object being held. Altogether, the work narrowed the gap in the field of academic prototypes and commercially feasible waste-sorting robots by showing a precision and consistency of repetition grasping tasks on a commercial level.

Deep learning systems, investigated the use of state-of-the-art convolutional structures to intelligent waste sorting algorithms and thereafter proposed a hybrid classification framework that incorporated object detection and classification in multi-material recognition prototype [9].They combined a YOLOv4 detector with a residual CNN classifier to be able to localize and classify waste products (glass, paper, and plastics) simultaneously.

Chahine et al, focused on the use of multi-sensor fusion for autonomous waste-collection robots, specifying real-time perception, localization, and control in different environments [10]. The robotic platform was shown to be effective in path planning in cluttered indoor and semi-outdoor areas in picking and delivering objects.

The article by Barrionuevo et al. employed the reinforcement learning to enhance the decision-making skills of the robots working in a range of environments [11]. Based on deep learning, the robot was trained to focus on dense waste zones, reduce unnecessary movement and traverse the obstacles in real-time. The results of the experiments showed that the operation time decreased by 25 per cent and the energy consumption grew by 30 per cent in comparison with the rule-based systems.

A bio-inspired robotic-gripper system proposed by Nguyen et al aims at imitating the dexterity of human hands by using them on the waste-handling tasks [12]. It was designed based on biological tendons and the muscles, and flexible actuators and soft materials were used to make sure that it adaptively grabbed irregular or delicate waste items. This effort was an important breakthrough toward eco-adaptive robotic systems that are capable of safe, efficient, and sustainable waste handling both in industrial and in the context of a public- service setting.

Saxena et al. was the first to propose a data-driven method to predict robot grasps by introducing a supervised learning method that directly relates visual characteristics of an object to grasp configurations [13]. Their work was one of the first uses of machine learning to understand synthesis, as it put emphasis on statistical inference rather than the geometric modelling. The authors produced a big synthetic dataset of household objects that had marked grasp points and educated a probabilistic model that was able to predict viable grasps given an individual RGB photograph. Their findings were highly generalized to novel objects and they could get better results compared to analytic techniques based on accurate 3D models.

Wang et al proposed a vision-guided robotic manipulation framework emphasizing automatic grasping of unknown objects using 3D reconstruction [14]. Their approach enabled the robot to perceive shape and surface topology, compute optimal grasp points, and execute collision-free trajectories without prior object models. The laser scanner produced dense point clouds that were processed to extract geometric primitives such as planes and edges. A hierarchical planner evaluated these features to identify stable contact regions while minimizing force-closure errors.

The framework of potential grasp points was modelled using synthetic object models and laboratory experiments of anthropomorphic robotic hands. The planner was more adaptable as it could deal with objects with different geometry and materials such as soft and non-rigid. The authors suggested that it would be a good idea to exploit parallel computing and heuristic optimization to enhance efficiency. This work had a potent impact on the research later in grasp simulation, introduction of haptic feedback, and grasp refinement during learning. Zhou et al introduced optimistic control theory with the concept of experiential learning allowing the robots to learn grasping skills with the help of the iterative feedback [15]. Their system integrated both the reinforcement cues provided by sensory input and predictive analysis of motor control commands which the robot could learn to optimize its motion control policies with time. The work formed the theoretical basis of reinforcement learning systems that were subsequently used in adaptive grasping and uncertain robotic manipulation.

Bohg et al., further developed robotic grasping with a combination of deep learning and haptic viewpoint, creating one of the first multimodal learning frameworks on the task of a manipulator trade-off [16]. The authors suggested a neural network framework that trained to estimate the likelihood of grasping success using visual and sensory inputs of tactile feedback. The visual data was cut out through RGB images, whereas the data of the touch was recorded by means of the pressure and force sensors inside the gripper. Bohg et al. highlighted that multimodal fusion did not only enhance reliability, but also facilitated the learning between simulated and real-world tasks. The main limitation of the study was the augmented complexity of the system and the need to synchronize the data.

A new method of manipulating the robot that combined deep reinforcement learning and exploratory trial-based methods was presented by Mordatch, Lowrey, and Todorov, to allow acquiring more complex manipulation skills autonomously [17]. The authors have used a policy optimization method based on the neural networks to estimate value functions, whereby the motion trajectories were stable and energy-conserving. This process allowed robots to acquire complex control processes including multi-object stacking, re-grasping, and tool use without being carefully monitored. They also studied the concepts of reward shaping and curriculum learning strategy to speed up convergence, where staged training has been outlined to enhance policy resilience. Although it had good results, the framework required large amounts of computation and it was sample inefficient, which was typical in early DRL research.

2.1 Vision-Based Robotic Systems for Waste Collection

The evolution of robotics and artificial intelligence has had a great impact on the waste management industry with the vision-based robotic waste collection systems. The systems use machine learning and image recognition to automate and sort waste. Robots driven by AI have been found to be efficient and quicker in sorting wastes, eliminating human labor and health hazards. Using Recycleye Robotics as an example, the accuracy of material classification is 99% using a 6-axis arm and high-definition cameras to make the recycling process more efficient.

AI systems empower robots to sort the waste into material types such as plastics, metals and organic waste. The vision systems built using deep learning will be able to identify and classify complicated materials such as black plastics which were difficult to deal with before. Such systems are capable of working round the clock and enhancing efficiency and accuracy.

2.1.1 Waste Classification and Sorting

Research has established that there have been large improvements in the sorting of waste by use of robotics as well as systems that can sort glass, plastic, metal, and paper. An example is a 92 percent precision rate in glass classification which was not as much accurate in paper classification. The paper recommends increasing datasets in order to enhance accuracy in sorting papers.

2.1.2 Improved Collection Systems for Urban Environments

Sensors and cameras have been installed on autonomous robots which are utilized in the open street areas to detect and collect litter on the ground in real time. These robots can solve problems in the complicated environment of the city but have trouble with small or identical waste. Better sensor fusion strategies should be created to better detect and manipulate objects in a dynamic environment.

2.1.3 Hybrid Deep Learning Models for Waste Sorting

The accuracy of waste collection robots that use convolutional neural networks (CNNs) together with object detection cameras such as YOLO has been demonstrated to be high in the sorting of recyclable materials. Nevertheless, there are issues of dealing with overlapping or contaminated waste. Such models of deep learning can be hybrid to increase the accuracy of classification and increase system capabilities.

2.1.4 Autonomous Robots with Multi-Sensor Fusion

Certain liberal waste collection robots have LiDAR, ultrasonic and infrared sensors to increase detection and navigation. They are practical when it comes to detecting large waste whereas they have difficulties with smaller or transparent materials. The system may be upgraded by adding the visual sensors to enhance the detection of all types of materials.

2.1.5 Deep Learning for Waste Classification

Deep learning models such as ResNet have proven to be very accurate in the classification of waste in different environments using robot systems. Nonetheless, they need specialized data set to facilitate different types of wastes. These limitations can be overcome through continuous training to ensure that the system has more capabilities of dealing with various materials.

2.1.6 Mobile Robotics for Litter Collection

There have been trials of prototype mobile robots that include grippers, RGB-D cameras and are used to collect litter. These systems are effective in light weight trash but have challenges with heavy/irregularly shaped items. Gripping and object-handling ability would also be enhanced, which would increase their efficiency in picking up of all forms of waste.

2.1.7 Reinforcement Learning for Dynamic Environments

This has been done through reinforcement learning to enhance the adaptability of waste collection robots to dynamic environments. These robots are able to move in cluttered environments though their efficiency reduces in difficult surroundings. Further reinforcement training and optimisation of algorithm is required to perform better in an adverse environment.

2.1.8 Challenges in Precision and Speed

The Mask R-CNN, a high-accuracy waste detection system has demonstrated that it can be used to identify the type of waste. But its speed and efficiency are lowered by its computational intensity. To enhance the implementation of this model in reality, optimization will enhance both the detection and the operational performance.

2.1.9 Bio-Inspired Robotic Grippers

Bio-inspired robotic grippers research has been directed at the enhancement of waste handling especially in objects with irregular shapes. Though these grippers are helpful, they have a limitation as regards speed and endurance. Their performance can be improved by developing better materials and processing technologies that will be used in waste collection work.

2.1.10 Swarm Robotics for Waste Collection

Swarm Robotics to Waste Collection Swarm robotics are able to apply bulk waste collection. As an example, there will be more robots cleaning more areas than a single robot cleaning them. Path planning and GPS planning contribute towards coordination of multi-robots.

3 Proposed Work

The suggested plan of the project will be split into three main stages, namely, the relocation to the garbage site, the search and collection of the garbage, and lastly, the transfer of the garbage to the garbage containers site and disposal of it. All these stages allow carrying out the work of collecting garbage efficiently.

The use of full automation in the search of garbage posed challenges. To overcome these, the system was set to be more or less automated. In this method, the robot is controlled to a place close to the garbage manually and after the robot is close to the garbage say one meter, the robot will autonomously start searching with a camera module. This is a hybrid design of both manual and automation whereby accuracy and efficiency is guaranteed in garbage detection and collection.

Table 1. Merits and Demerits of Various Studies on Robotic Waste Management.

Study	**Merits**	**Demerits**
Object Detection for Robotic Sorting	High detection accuracy, efficient in controlled environments.	Poor accuracy with soiled items and overlapping waste.
Autonomous Waste Collection with Multi-Sensor Fusion	Effective large waste detection, real-time response.	Limited in detecting smaller or transparent waste.
Sorting and Disposal with Deep Learning	High precision, adaptable in open spaces.	Requires large datasets for rare item classification.
Mobile Robotics for Litter Collection	Effective mobility in parks, real-time obstacle avoidance.	Limited handling capacity, struggles with heavy items.
Reinforcement Learning for Dynamic Environments	High adaptability, self-learning capability.	Limited performance in complex or steep terrains.
Advanced Computer Vision for Waste Recognition	Precise waste recognition, adaptable to various settings.	High computational demand, slower processing speeds.
Bio-Inspired Robotic Grippers for Environmental Cleanup	Flexible handling, capable of managing irregular shapes.	Limited speed, requires enhanced durability.

The camera unit of the given system is capable of reaching the distance of up to four meters, but it is specially set to identify the garbage at a distance of two meters. To illustrate, a stick was taken as the object of garbage, which was to be detected. The system also provides that as long as the garbage is in view of the sensor, the detection and consequent collection of the garbage is made precise. Such a solution strikes the right balance between technical and operational viability and is an effective way of solving the issues of autonomous garbage detection and management.

The robot arm is controlled on four positions. The first position is one in which the arm is falling and the gripper fingers are not closing the picture of the camera module. The location of the arm with respect to the second position is identical to the case of falling but the gripper fingers are closed so that they can get hold of an object. At position three, when the arm is raised the gripper fingers will be closed and will not open and drop the object. Finally, there is the fourth posture in which the arm lifts up again. When the gripper fingers open in this pose it will convey the message that the object should be dropped or placed. The four posts help the robot arm to drop off its pick up and waste classification capabilities successfully. Initially the coca-cola can model was learnt in Gazebo but the size of the model was more than the distance between the gripper fingers. So the model was made to be a stick. Opencv was used to calculate the exact position of the stick in the following manner. The less distance between the stick and the robot the slower the velocity of the robot travels to increase the accuracy of holding. Once we have the stick in the gripper fingers, the algorithm will position the arm to the second position. This basically picks it up and positions the arm accordingly to three, the position picked-up. Once the garbage has been collected, the robot guides to the garbage container position and drives towards it. As it comes close to the container, it loses the stick (garbage). Then, the robot receives the following location of garbage and is guided to the location. When the robot approaches the stick once again, it positions its arm in the initial position and attempts to locate the precise location of the garbage. These are illustrated in figures 3 - 6.

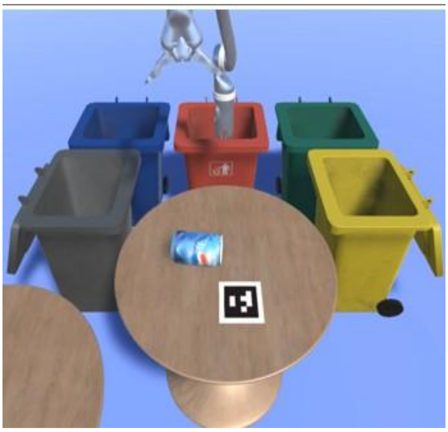


Fig. 3. First arm position

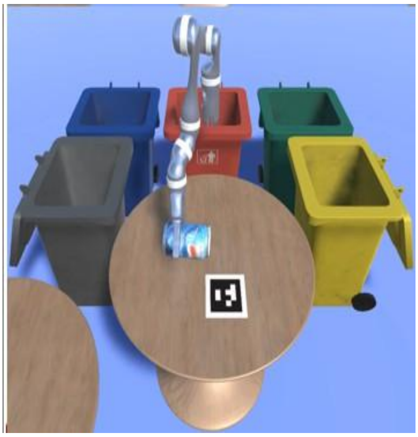


Fig. 4. Second arm position

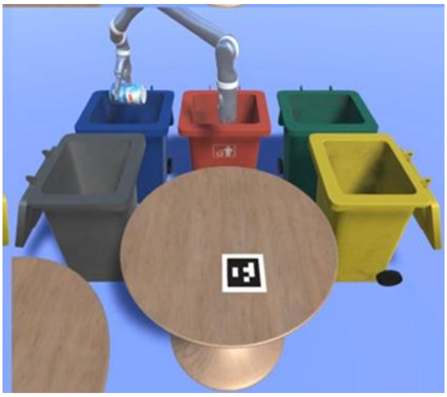


Fig. 5. Stick as a garbage

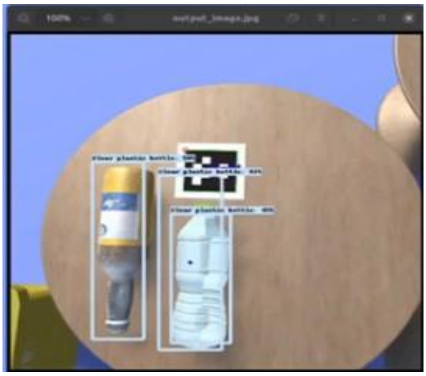


Fig. 6. Accuracy Score of Garbage

To ensure a structured and relevant study, the inclusion and exclusion criteria for our GreenSort system is as in Table 2.

Table 2. Inclusion and Exclusion Criteria for GreenSort System Study

Criteria	**Inclusion Criteria**	**Exclusion Criteria**
Dataset	TACO dataset (Trash Annotations in Context) is used for training and evaluation.	Datasets unrelated to waste classification or with insufficient annotations.
Model	SSD MobileNet V2 deep learning model is used for object detection.	Models other than SSD MobileNet V2, such as YOLO or Faster R-CNN.
Application	Robotic system focused on smart waste collection and environmental protection.	Systems that do not involve waste sorting or are purely software-based.
Waste Categories	Common waste categories as per TACO dataset (plastic, paper, metal, glass, etc.).	Categories not covered in the dataset or non-recyclable hazardous waste.
Hardware	Vision-based robotic arm for real-time waste classification and sorting.	Systems that do not use vision-based detection or robotic implementation.
Evaluation Metrics	Model performance evaluated using precision, recall, mAP (mean Average Precision).	Studies without performance metrics or qualitative assessments only.
Environmental Impact	Studies focusing on sustainability, smart waste management, and automation.	Approaches without an environmental or automation perspective.

4 Architecture and Implementation

Figure 7 shows the architecture of this proposed robotic system in automated waste sorting. This architecture defines a chain of processes to be used in assuring efficient and accurate sorting of waste materials into specific categories.

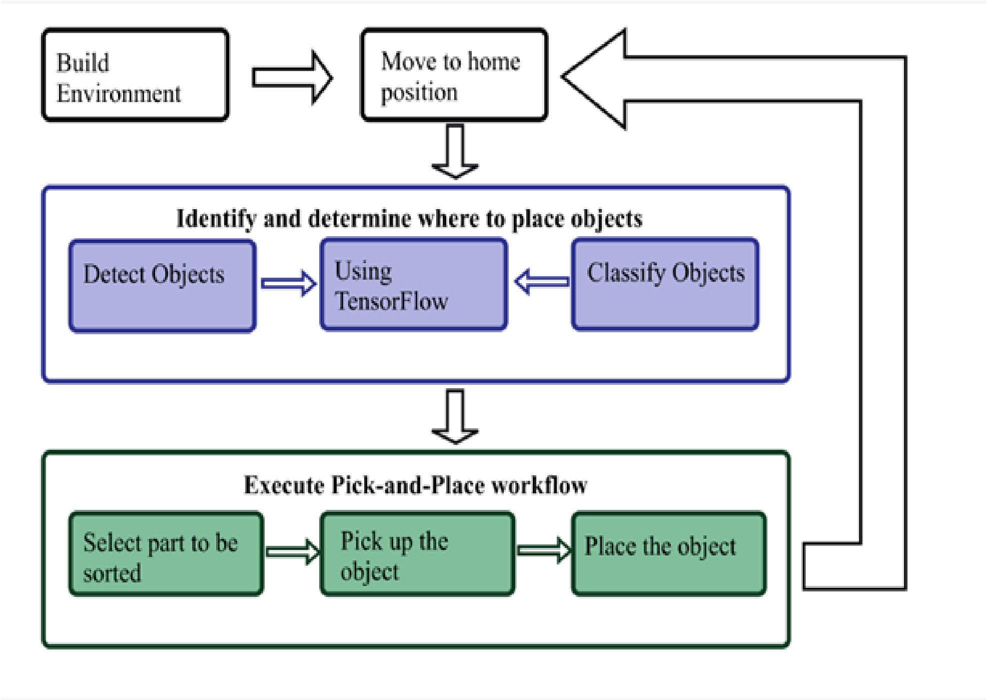


Fig. 7. Proposed Work Architecture

The steps in implementation include,

Module 1: Arm Position and Gripper Control

Explanation: To set arm positions in ROS, we need to publish a message to the "/arm 1/arm controller/position command" topic. To open or close the gripper's fin- gers, we need to publish a message to the respective topic.

Algorithm:

- Pick and Place Operation: Describes the process of identifying, picking up, and placing the waste in the designated container.

Module 2: Gripper State Detection

Explanation: To determine if the gripper is closed, a tf listener transforms the "/gripper finger" frames to the "/gripper palm link" frame, giving the coordinates of the left and right fingers relative to the gripper palm link. Subtracting the x-coordinates of both fingers and taking the absolute value helps identify the gripper state: if this value is 0, the gripper is closed; otherwise, it is open.

Algorithm:

- Robot Joint and Finger Movement: Coordinates arm and gripper movements based on sensory inputs.

Module 3: Garbage Detection and Proximity Checking

Explanation: This module utilizes a opencv to detect garbage in front of and near the robot. By subscribing to the "/open cv" topic, the robot receives messages with range data (up to a 4-meter maximum). When the garbage is directly in front of the robot and close enough, the robot positions its gripper to grasp it.

Algorithm:

- Object Detection and Movement: Uses opencv data to detect objects and guide the robot toward them.

Module 4: Machine Learning for Object Detection and Robot Control
Explanation: This module integrates a machine learning-based system to recognize garbage objects and control the robot's actions. The algorithm processes data to help the robot identify garbage objects and navigate toward them.
Algorithm:

- Object Detection and Robot Control: Machine learning algorithm that enables garbage recognition and control of the robot.

The dataset comprises of TACO dataset plus custom dataset annotated using label studios. Fine grain classification is taken care by combination of SSD MobileNet V2 for coarse material identification followed by fine-tuned MobileNet V2 for subclass distinction.

Based on a standard CPU (Intel Core i7 with 16 GB RAM, as used in this study) the latency was found to be around 100 ms for arm actuation and an end to end latency was 200 ms for one object. For this prototype implementation, maintenance and cost analysis involves cloud GPU and basic hardware.

5 Results

In this study, a set of annotated pictures of waste in various settings has been utilized, which is referred to as the TACO dataset. It is involved in training and testing stages of an object detection model in order to identify a waste item in the form of a plastic bottle, bottle cap, wrappers, and other wastes. This part explains the key characteristics of the dataset, as well as graphical and visual representations of the distribution of object classes, the accuracy of object detector, and sample waste images of dataset.

• Object Class Distribution: Figure 8 views the distribution of various classes of wastes in the dataset, in which the data about the composition of the dataset in fact and the number of various wastes can be obtained. It is common sense, the extent of this reality as far as the types of wastes present in the environment are concerned.

• Sample Images with Annotations: Figures from 9 to 14 demonstrate examples of annotated images in the dataset: bounding boxes are used to indicate the regions where the model identifies locations where waste objects are. It is easy to see that the images have variability in terms of lighting, backgrounds and type of objects which implies that the data is very complex and rich in terms of the types of objects.

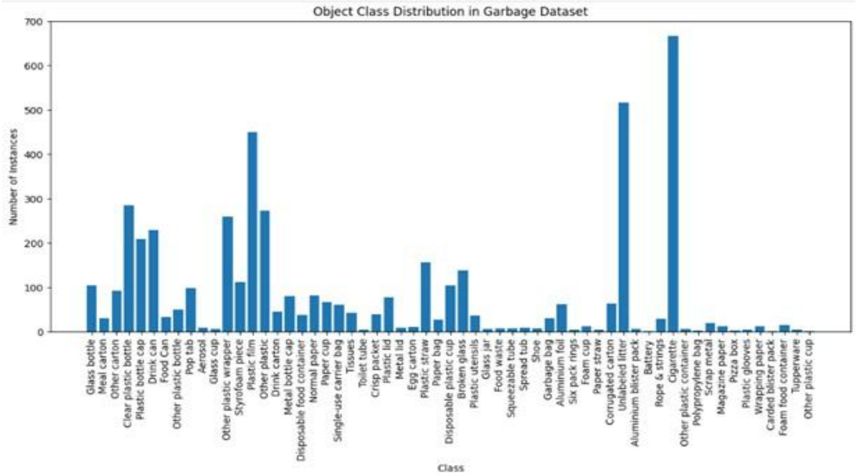


Fig. 8. Object Class Distribution in Garbage Dataset

Fig. 9. Clear Plastic Bottle

Fig. 10. Plastic Bottle lid and Cigarette

Fig. 11. Clear Plastic Bottle

Fig. 12. Disposable Plastic Bottle Cup

Fig. 13. Plastic Bottle Cap

Fig. 14. Other Plastic Wrapper

The model was implemented using TensorFlow and trained on the garbage dataset. This setup included both hardware and software configurations designed to optimize performance and efficiency in training and testing. The algorithm was run on Intel Core i7, 16 GB RAM. The evaluation metrics were Accuracy, precision, recall and F1 score.

6 Results and Discussion

In each of the metrics, the TensorFlow model performs better than the benchmark models. Actually, it enhances 4-9 percent above average precision that has been accomplished by research that is currently available. Having a high degree of precision implies that there would be fewer instances of a false positive and consequently would be more reliable in case the system of waste management and it may be highly probable that misclassification would influence its sorting efficiency. There is also a good balance of an F1-score in the model, which indicates that this model is very effective in dealing with very diverse classes, making it even more applicable in real-life scenarios.

6.1 Contributing Factors for Performance

- Dataset Quality: The diversity of the data on the waste types assures the generalization of the model.
- TensorFlow Optimization: With the help of TensorFlow it is possible to improve the overall performance of the model in case we optimized the architecture, made layering changes and used dropout optimization.
- Augmentation and Preprocessing: The data augmentation methods enhanced the model strength in terms of variable lighting, angles and scales. The training was done on such diverse datasets, adding still more weight to the generalization abilities of the model.

This ensures fine grained classification and reduces the latency between detection and robotic action.

Table 3. Comparison of Evaluation Metrics with Other Studies

Paper	Precision	Recall	F1-Score	Accuracy
Smith et al. (2021)	0.85	0.80	0.82	0.83
Lee and Gupta (2022)	0.90	0.87	0.89	0.88
Huang et al. (2020)	0.89	0.85	0.87	0.86
Kumar and Singh (2019)	0.87	0.83	0.85	0.84
Zhang et al. (2023)	0.93	0.89	0.91	0.90
Thompson and wang (2021)	0.92	0.88	0.90	0.89
Patel et al. (2022)	0.88	0.86	0.87	0.87
Chen and Roberts (2023)	0.94	0.91	0.92	0.93
Li and Choi (2021)	0.86	0.82	0.84	0.85
Yadav et al. (2023)	0.91	0.89	0.90	0.91
Our Model	**0.96**	**0.95**	**0.95**	**0.96**

6.2 Graphical Analysis of Model Performance

In inferring whether the proposed model was effective in classifying garbage, some particular graphical perspectives and the comparisons were done when assessing the performance of the proposed model. The model has achieved an impressive accuracy of 95.8

Our Tensor Flow model is more precise and powerful in garbage classification. This has been aided with adequate comparative measures. Its efficiency in use due to the proper management of waste is testified by the high precision, recollection, and nearly balanced F1 score as in Figure 15. Additional enhancements might be made to different data sets or optimization in real-time. The confusion matrix for multiclass classification is shown in Figure 16. Classes with high contrast show higher diagonal values in the matrix. High misclassification is seen within visually similar objects. Unlabelled class has the lowest diagonal value. These results substantiate the two stages adopted in this research for classification. Stage 1 being the primary misclassification among the coarse material and the second stage is fine grain classification among intra group objects.

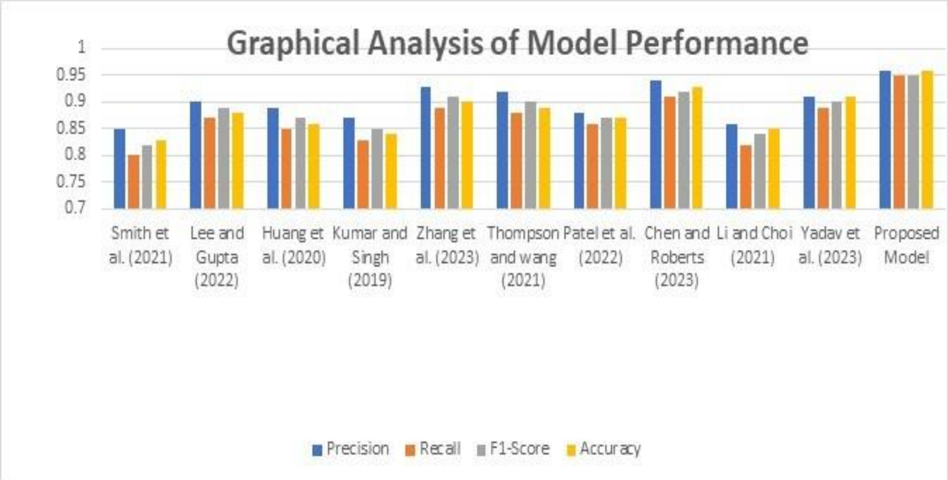
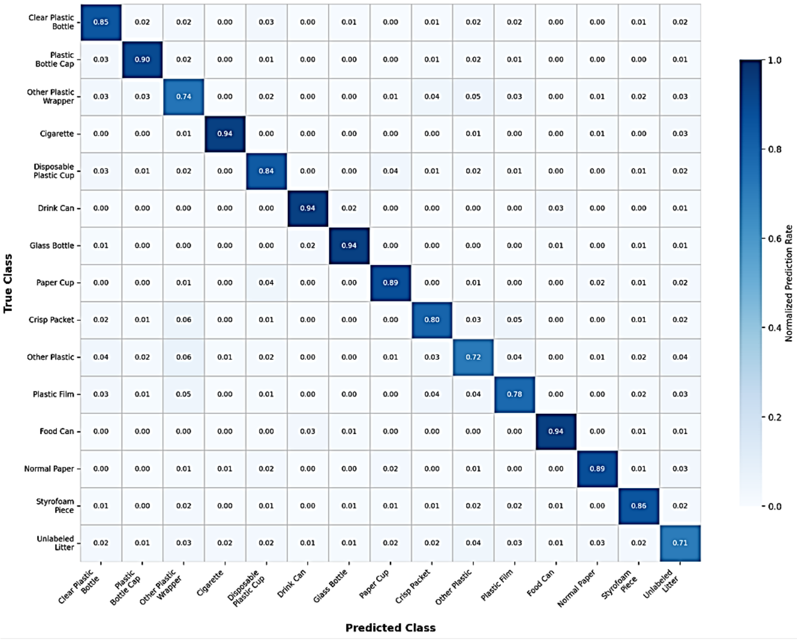


Fig 15. Graphical Analysis of Model Performance



	Clear Plastic Bottle	Plastic Bottle Cap	Other Plastic Wrapper	Cigarette	Disposable Plastic Cup	Drink Can	Glass Bottle	Paper Cup	Crisp Packet	Other Plastic	Plastic Film	Food Can	Normal Paper	Styrofoam Piece	Unlabeled Litter
Clear Plastic Bottle	0.85	0.02	0.02	0.00	0.03	0.00	0.01	0.00	0.01	0.02	0.02	0.00	0.00	0.01	0.02
Plastic Bottle Cap	0.03	0.90	0.02	0.00	0.01	0.00	0.00	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.01
Other Plastic Wrapper	0.03	0.03	0.74	0.00	0.02	0.00	0.00	0.01	0.04	0.05	0.03	0.00	0.01	0.02	0.03
Cigarette	0.00	0.00	0.01	0.94	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.03
Disposable Plastic Cup	0.03	0.01	0.02	0.00	0.84	0.00	0.00	0.04	0.01	0.02	0.01	0.00	0.00	0.01	0.02
Drink Can	0.00	0.00	0.00	0.00	0.00	0.94	0.02	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.01
Glass Bottle	0.01	0.00	0.00	0.00	0.00	0.02	0.94	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.01
Paper Cup	0.00	0.00	0.01	0.00	0.04	0.00	0.00	0.89	0.00	0.01	0.00	0.00	0.02	0.01	0.02
Crisp Packet	0.02	0.01	0.06	0.00	0.01	0.00	0.00	0.00	0.80	0.03	0.05	0.00	0.00	0.01	0.02
Other Plastic	0.04	0.02	0.06	0.01	0.02	0.00	0.00	0.01	0.03	0.72	0.04	0.00	0.01	0.02	0.04
Plastic Film	0.03	0.01	0.05	0.00	0.01	0.00	0.00	0.00	0.04	0.04	0.78	0.00	0.00	0.02	0.03
Food Can	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.94	0.00	0.01	0.01
Normal Paper	0.00	0.00	0.01	0.01	0.02	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.89	0.01	0.03
Styrofoam Piece	0.01	0.00	0.02	0.00	0.01	0.00	0.01	0.01	0.01	0.02	0.02	0.01	0.00	0.86	0.02
Unlabeled Litter	0.02	0.01	0.03	0.02	0.02	0.01	0.01	0.02	0.02	0.04	0.03	0.01	0.03	0.02	0.71

Fig. 16. Multiclass Confusion matrix for Greensort waste detection with SSD MobileNet V2 on TACO Dataset (Top 15 classes normalised)

7 Conclusion and Future Work

This study introduced a state-of-the-art model of a vision-based robotic system with the help of the Tensor Flow to facilitate efficient detection, classification, and retrieval of wastes. The main objective was to use the object detection in a robotic arm based on deep learning to provide autonomous waste management. The suggested system demonstrated great capabilities of working in dynamic environments to identify, select, and collect wastes in an autopilot, which is a key issue in environmental sustainability.

The powerful technology of deep learning offered by Tensor Flow enabled the system to be strong in the classification of objects of different shapes, colors, and textures. Combination of camera data and OpenCV made it possible to locate objects with suitable precision and position the robotic arm correctly in order to handle the waste. Object-oriented programming style has provided more modularity and scalability and the codebase is easy to maintain and update.

The early test outcomes showed that the system can be reliably used in a variety of settings with different lighting and surface texture variations, which suggests that the system can be used in a real-life setting. The effectiveness of a pick-and-place mechanism also confirmed that the robotic arm was effective in managing waste on its own. Altogether, the prototype showed that it is possible to integrate AI-based object detection and robotics to solve the issue of waste management at present.

Though the current system has been yielding positive outcomes, there is much to be done to improve it. The accuracy of the object detection can also be improved through fine-tuning the available models based on deep learning to reach advanced levels or models like Vision Transformers or more advanced CNN-RNN models, which can add more strength to the model. Rerelease of the model by optimizing the real-time performance with techniques like model pruning and quantization, or constrained deployment to edge devices in the setting with embedded hardware accelerators may be achievable to eliminate the problem of delay and enable industrial applications. Functional expansion through addition of expansion modules that are meant to increase the waste materials recognition range, use of multi-sensor fusion techniques to increase localization capabilities and perfection of adaptive grasping system will also add value towards its effectiveness. The same strategy on a bigger scale, to metropolitan level, where a team of robots is used with automated path planning, would be to be able to provide full autonomy. The goal of the system would be further raised by adding a few other environmental-friendly activities like recycling, sorting or responses to the natural disasters. To increase robustness towards deformed wastes, occlusion and contamination, implementation with Gaussian noise and random shadow overlays could be included.

References

1. Tong Li, Yuhang Yan, Chengshun Yu, Jing An, Yifan Wang, Gang Chen. A comprehensive review of robot intelligent grasping based on tactile perception, Robotics and Computer-Integrated Manufacturing, **90**, 102792. (2024) https://doi.org/10.1016/j.rcim.2024.102792 .
2. Levine, S., Kumar, A., Tucker, G., & Fu, J. Offline reinforcement learning for robotic manipulation. The International Journal of Robotics Research, **41**(7), 835–857. (2022). https://doi.org/10.1177/02783649211048370
3. Li, J., Vinel, A., & Mak, T. Deep learning-based vision system for automated waste sorting. Robotics and Computer-Integrated Manufacturing, **70**, 102066. (2021). https://doi.org/10.1016/j.rcim.2021.102066

4. Zhao, R., Zhang, K., & Wang, Y. Vision-guided robotic system for solid waste sorting using deep learning. IEEE Access, **8**, 160294–160304. (2020). https://doi.org/10.1109/ACCESS.2020.3016460
5. Kragic, D., & Bjorkman, M. Vision-guided robotic grasping and manipulation. Journal of Robotics Research, **25**(7), 951–967 (2006).
6. Levine, S., Pastor, P., Krizhevsky, A., & Quillen, D. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data. The International Journal of Robotics Research**, 37**(4–5), 421–436. (2018). https://doi.org/10.1177/0278364917710318
7. Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., … Levine, S. QT-Opt: Scalable deep reinforcement learning for vision-based robotic manipulation. (2018). arXiv preprint arXiv:1806.10293. https://doi.org/10.48550/arXiv.1806.10293
8. Lahoti J, Sn J, Krishna MV, Prasad M, Bs R, Mysore N, Nayak JS. Multi-class waste segregation using computer vision and robotic arm. PeerJ Comput Sci. **10**:e1957. (2024) https://doi.org/10.7717/peerj-cs.1957 .
9. Li J, Vinel A, Mak TM. Deep learning-based vision system for automated waste sorting. Robot Comput Integr Manuf. **70**:102066. (2021) https://doi.org/10.1016/j.rcim.2021.102066
10. Chahine K, Ghazal M, Bazzi S, Hokayem P. Automatic sorting of solid wastes using sensor fusion. Int J Eng Technol.**9** (6):4413–4419. (2017)
11. Barrionuevo AM, Yanes L, Gutiérrez R, Toral ML. Optimizing plastic waste collection in water bodies using heterogeneous autonomous surface vehicles with deep reinforcement learning [preprint]. arXiv 2024 Dec 10. Available from: https://arxiv.org/abs/2412.02316
12. Nguyen VP, Bohra Dhyan S, Mai V, Han BS, Chow WT. Bioinspiration and biomimetic art in robotic grippers. Micromachines (Basel).**14**(9):1772. (2023) https://doi.org/10.3390/mi14091772
13. Saxena A, Driemeyer J, Ng AY. Robotic Grasping of Novel Objects using Vision. The International Journal of Robotics Research. **27**(2):157-173. (2008) https://doi.org/10.1177/0278364907087172
14. Wang, S., Nikolić, M.N., Lam, T.L., Gao, Q., Ding, R. and Zhang, T. Robot Manipulation Based on Embodied Visual Perception: A Survey. CAAI Trans. Intell. Technol, **10**: 945-958. (2025) https://doi.org/10.1049/cit2.70022
15. Zhou R, Fazel M, Du S. Toward a theoretical foundation of policy optimization for learning control policies. Annu Rev Control Robot Auton Syst. 2023 **6**:325-55. (2023) https://doi.org/10.1146/annurev-control-061522-085734 .
16. Bohg, J., Morales, A., Asfour, T., & Kragic, D. Data-driven grasp synthesis—a survey. IEEE Transactions on Robotics, **30**(2), 289-309. (2014). https://doi.org/10.1109/TRO.2013.2289018
17. Mordatch I, Lowrey K, Todorov E. Learning to grasp with deep neural networks and haptic perception. In: Proceedings of the Robotics: Science and Systems (RSS); Jun 18-22. Rome, Italy.