International Journal of Mechanical and Electrical Engineering
ISSN: 3005-9615 (Print), ISSN: 3005-7132 (Online)
Volume 8, Number 6, Year 2026
DOI: https://doi.org/10.62051/ijmee.v8n6.01

Design of an Intelligent Waste-Sorting Robot Based on Image Recognition

Changkun Luo

College of Engineering, The University of California, Santa Barbara, CA 93117, USA

ABSTRACT

The growing waste crisis in the world at large requires automated sorting to improve on recycling efficiency and advance the application of the circular economy. In the paper, the developed intelligent waste-sorting robot combines the deep learning-based on YOLOv8 with robot manipulation systems to classify waste and sort it automatically. The given system is implemented with the usage of an RGB-D camera to obtain images, noise reduction and normalization preprocessing algorithms, and a YOLOv8 neural network to detect multiple classes of objects in 9 categories of waste. A 6-DOF collaborative robotic arm with adaptive gripper performs physical sorting tasks computed on the basis of the 3D coordinates and motion planning algorithms. Using an experimental dataset of 8,500 annotated images, the 94.3% mAP at 0.5 detect rate, 92.5% precision, 91.1% recall, and 8.7 ms inference time are achieved, which makes it possible to process images at 115 frames per second in real-time. Robotic manipulation module has 92.3 percent grasping success and the system throughput is 445 objects/hour that is impressive when compared to manual sorting techniques. The findings confirm the feasibility of the system in practice in two areas: municipal recycling centers and industrial waste management and also in niche applications such as e-waste processing.

KEYWORDS

YOLOv8; Object Detection; Recycling Automation; Convolutional Neural Networks; Sustainable Waste Management.

1. INTRODUCTION

Waste crisis is one of the most acute environmental issues of the 21st century that is increasing all over the globe. It is estimated that presently the world has municipal solid waste production of about 2.01 billion tons every year and it is projected that the figure will reach 3.40 billion tons in 2050. The main cause of this dramatic increase in waste generation is the fast urbanization, population increase and industrial development which have complex implications that do not only limit to land pollution but also to groundwater pollution, emission of greenhouse gases, loss of ecosystem, and posing serious public health risks.

Conventional management methods of waste, especially manual sorting exercises, have become unsatisfactory to handle the magnitude and the complexity of the present waste patterns. Not only are the manual sorting activities laborious and time-consuming but also exposes the people performing the work to risky conditions such as biological hazards, chemical exposure as well as ergonomic-related injuries. Research has established that about 90 percent of the garbage produced in most developing areas ends up in landfills without any sorting procedures thus causing a lot of environmental damages and waste of resources. This is made possible by the fact that the efficiency

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

of traditional technologies is extremely low (possibly no more than 70-80 percent of the sorting accuracy) and the necessity to develop new technologies in waste management.

Combination of artificial intelligence, computer vision and robotics has provided unprecedented opportunities to revolutionize the practice of waste management. Convolutional Neural Networks (CNNs) and object detection algorithms, including YOLO (You Only Look Once), are deep learning technologies that have proved to be amazing in waste identification and classification problems[1]. From recent studies, it has been demonstrated that AI-based waste sorting devices can reach detection rates of up to 85-99 percent, which is much higher than the ability of humans to detect waste and much higher than that of humans because sorting is a 24/7 task [2]. A combination of these technologies, together with robotic manipulation systems, allows automated real-time sorting of waste at high rates and at levels of precision that are not achievable by the manual process.

Visual classification refers to the use of image recognition as a way of waste sorting as opposed to old spectroscopic waste sorting techniques. In the modern world, advanced deep learning models such as YOLOv8, ResNet, MobileNet, and DenseNet are used, and they have been specifically adapted to waste detection tasks [3]. The models use transfer learning, attention-based models, and multi-scale feature extraction to address the intrinsic weaknesses of waste classification such as the varieties of shapes of objects, different light conditions, and various types of occlusions, contaminations, and material degradations [4]. The market of robotic waste sorting has increased significantly, reaching USD 2.3 billion in 2023 and up to USD 3.4 billion in 2034 in valuations as automated sorting technologies are recognized more and more worldwide as important infrastructure in the circular economy.

Nevertheless, even with considerable advances, smart waste-sorting robots still have a variety of serious tasks that need to be resolved with the help of innovative solutions. The main issues are high classification accuracy of heterogeneous waste materials, strong performance in dynamic environmental conditions, low computational cost of real-time processing, reliable grasping of deformable and irregular objects, and economy of scale in large-scale application [5]. Also, the movement of the controlled laboratory conditions to the real-life working conditions is accompanied by significant technical challenges associated with system reliability, maintenance needs, and compatibility with the already existing waste management infrastructure.

This study attempts to overcome these difficulties by providing a detailed design architecture of an intelligent waste-sorting robot, which combines the latest image recognition methodology with the latest robotic manipulation structures. The proposed system architecture focuses on the real uses of the deployment such as the efficiency in computations, the capability of real-time processing, the accuracy in the classification, and reliability of the system. It combines recent innovations in the field of deep learning, computer vision, and robotics, which is why the current research is relevant to the creation of the next-generation automated waste management techniques that will be able to support the idea of sustainable resource recovery and the development of the circular economy.

2. LITERATURE REVIEW

In the recent past, there have been breakthroughs in the field of deep learning which have triggered seismic changes in automated waste sorting systems. The development of the old methods of computer vision to advanced neural network structures has made it possible to achieve new accuracy in the classification of wastes and the possibility of real-time detection. The modern studies aim at developing the best detection algorithms, improving the robotic manipulation system, and creating the unified systems that will connect the image recognition with the physical sorting tasks.

YOLO-based architectures have become the new standard of waste detection programs on a real-time basis. The system of garbage classification by Zhou et al. [6] on the basis of the YOLOv8 reached the accuracy of recognition of 90 percent at 35 frames per second, which proves the feasibility of its

implementation in unmanned vehicles. The paper emphasized that YOLOv8 had a better tradeoff between computational and detection accuracy than earlier. Likewise, Di et al. [7] also produced a model YOLOv8 with solid floating waste uniquely optimized features with improved feature extractors that outperformed YOLOv5, YOLOv9, YOLOv10, and YOLOv11 variants in terms of mean average precision and recall [12]. These results highlight the ongoing architectural improvements that allow YOLO models to deal with more complicated waste detection tasks.

The combination of ResNet and hybrid CNN networks has been proven to have a high level of performance. Jin et al. [8] came up with ResEff-YOLO that applies ResNet18d modules and Spatial Pyramid Pooling Fast (SPPF) modules alongside EfficientHead to enable multi-scale feature learning. The underlying problem that was solved by this architectural innovation was the detection of waste objects with extremely different sizes and orientations. The authors Mustapha et al. [9] proceeded with it and created a hybrid YOLOv8-CNN model, which combines the object localization capabilities of YOLOv8 and the classification capabilities of CNN to handle waste type classification in the real world with a higher accuracy degree, specifically distinguishing between compostable and non-compostable materials.

Special software has made AI-based waste sorting extend further than the usual municipal solid waste. Rajeev et al. [4] designed tailored YOLO models to e-waste classification, utilizing various backbone models that were EfficientNet, ResNet, and CSPDarknet to fine-tune the detection ability of electronic elements including batteries, circuit boards and cables. The paper has noted the relevance of the flexibility of the model to adapt to the various features of waste streams. Also, the analysis of floating waste localization in water has revealed the adaptability of deep learning methods, and improved YOLOv8 models have shown the capacity to work well in hard-to-detect scenarios in the outdoor environment with changing lighting and water surface reflections[10].

New studies consider more powerful segmentation methods other than bounding box detection. Vukicevic et al. [1] explored Segment Anything Model family to waste sorting in small-scale manufacturing setting as a versatile approach and compared SAM, FastSAM, MobileSAMv2 and EfficientSAM design. The two-step pipeline that they used to split the process of object extraction and classification proved to have greater flexibility with industrial applications where new forms of waste needed to be regularly adapted to. This method deals with serious constraints of specialized detection frameworks that do not scale to new waste mixture conditions.

One of the major research areas that are critical is level of integration in the system level. Recent literature is giving more focus to full robotic sorting solutions with computer vision, mechanical handling and data analytics. The implementation of the hyperspectral vision technology is quite an important development, and it has made it possible to distinguish between the materials that are visually indistinguishable, e.g. various types of plastic polymers or fabrics. Such capability overcomes one of the essential problems in waste sorting whereby the traditional RGB imaging cannot be used effectively. The future trends are indicated in the direction of large vision models that may be used to automatically annotate data and, as such, decrease the mandatory labeling of data and shorten the model development cycles.

Even though there have been significant improvements, there are still a number of research gaps. Existing technologies are largely controlled lighting and clean conditions, but in the real-world, recycling plants are extremely unstable with highly contaminated and partially obscured and deformed items. Future studies are needed on how to achieve resource-constrained hardware performance and robustness to new types of waste by combining these three attributes as a computational efficiency approach to edge deployment. The literature shows a trend of becoming more elaborate in terms of architecture but practical aspects of deployment such as cost-effective, maintenance and compatibility to the existing system are not adequately covered.

3. METHODOLOGY AND SYSTEM DESIGN

The following section contains the complete system architecture of the intelligent waste-sorting robot, which incorporates the image recognition algorithms with robotic manipulation systems. The suggested structure focuses on modularity, computational efficiency, and real-time processing to deal with real-life deployment needs.

3.1. System Architecture Overview

The smart waste-sorting robot is made of four main subsystems: (1) Image Acquisition Module, (2) Deep Learning Detection and Classification Module, (3) Robotic Manipulation Module and (4) Control and Communication Module[11]. The system takes the form of a conveyor belt system where the wastes are taken past the vision system to be identified and then sorted automatically using robotic gripping and placing in the relevant bins.

The general pipeline is designed in the form of a pipeline: the RGB-D cameras record the waste information on the conveyor belt and relay it to the processing unit, where the YOLOv8-based detection model defines the category of the objects and coordinates the bounding boxes. The control module computes the optimum grasp points depending on the depth data and object geometry and directs the robotic arm to perform pick and place operations [12]. The feedback loops are implemented in real-time to keep vision processing, motion planning and physical manipulation synchronized.

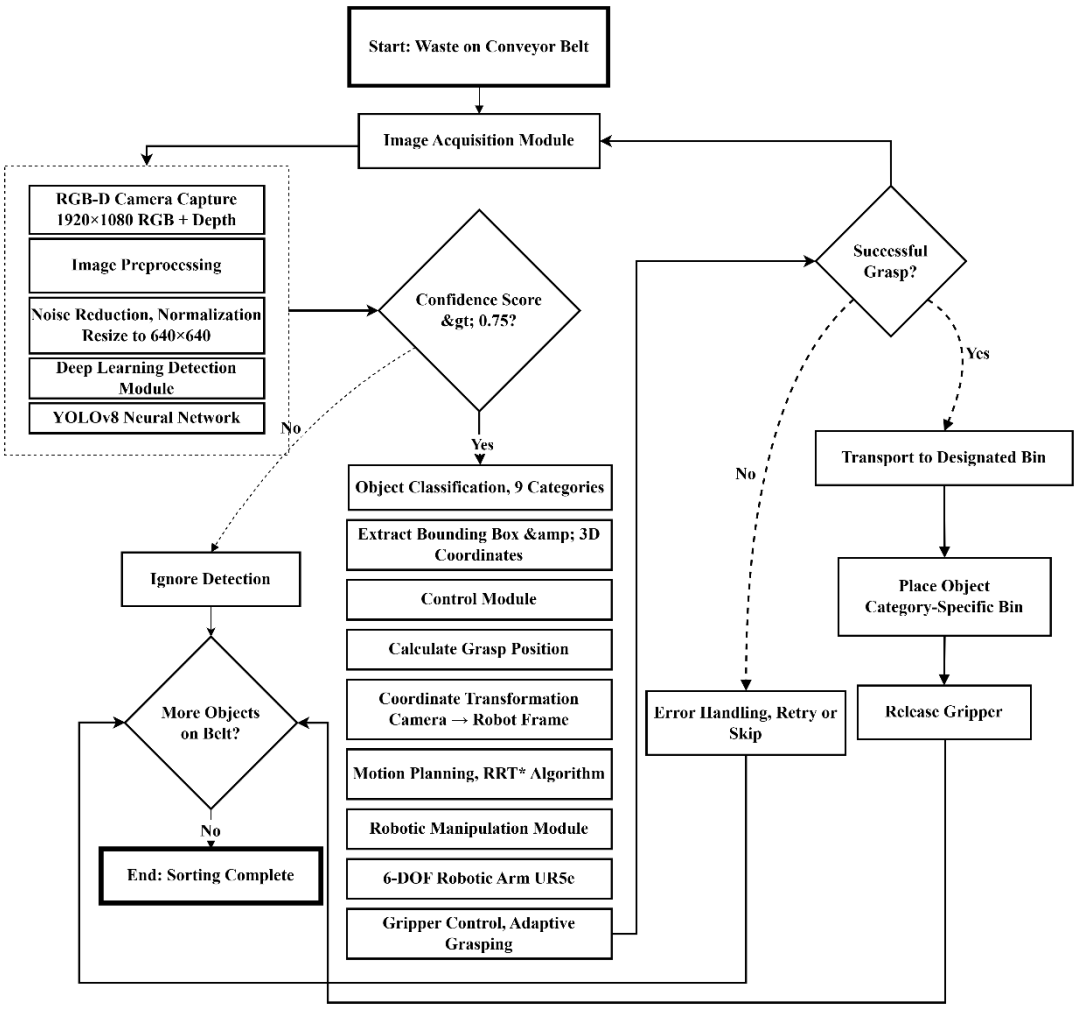

Fig 1. System architecture

3.2. Image Acquisition and Preprocessing

The camera used in the image acquisition module is the Intel RealSense D435 RGB-D which is positioned 0.8 meters above the conveyor belt and at an angle of 45 degrees to reduce the chances of occlusions and shadows. The camera has synchronized color (1920X1080 RGB) and depth (1280X720) frames with 30 frames per second. The raw images are preprocessed with various algorithms such as noise elimination or removal using the Gaussian filter, lighting normalization by using histogram equalization and isolating the waste objects on the conveyor surface using the background subtraction method [13].

Data augmentation techniques are applied during training to enhance model robustness: random horizontal flipping (probability 0.5), rotation (±15 degrees), brightness adjustment (±20%), contrast variation (0.8-1.2), and Gaussian noise injection ($\sigma = 0.01$). The preprocessing pipeline normalizes pixel values to the range [0,1] and resizes images to 640×640 pixels to match the YOLOv8 input requirements while maintaining aspect ratio through letterboxing.

$$Inorm = \frac{(I - \mu)}{\sigma}$$

3.3. Deep Learning Detection Model

The main detection system uses the YOLOv8, which is an anchor-free, single-stage object detector, optimized to operate in real time. YOLOv8 splits the input images into a grid and makes predictions of bounding boxes, objectness scores, and probability of classes directly without going through region proposal mechanisms. The architecture includes three categories of features: backbone (CSPDarknet53 to extract features), neck (Path Aggregation Network with Spatial Pyramid Pooling Fast to fuse features on a multi-scale), and head (decoupled detection heads to classify and perform regression).

The detection process can be formalized as follows. For an input image $I \in R$

$$P(class/object) \times P(object) \times IOU = Confidence\ Score$$

where $P(class|object)$ represents the conditional class probability, $P(object)$ denotes objectness confidence, and IoU measures the intersection over union between predicted and ground truth boxes. The model outputs predictions for C classes (plastic, paper, metal, glass, organic, hazardous, cardboard, textile, electronic) with bounding box coordinates ($x, y, w,$ h) representing center coordinates, width, and height.

Loss function optimization combines classification loss (binary cross-entropy), localization loss (Complete IoU loss), and objectness loss:

$$Ltotal = \lambda cls \cdot Lcls + \lambda box \cdot Lbox + \lambda obj \cdot Lobjb$$

3.4. Waste Classification Algorithm

The classification module implements a confidence-based decision mechanism with threshold filtering to eliminate false positives. Algorithm 1 presents the waste detection and classification procedure:

3.5. Robotic Manipulation System

The manipulation subsystem has a 6-DOF robotic arm (UR5e collaborative robot) with a pneumatic parallel-jaw gripper. Grasp point computation uses depth data to compute the 3D coordinates of the camera frame, and then transform to the robot base frame using hand-eye transformation matrices

that have been calibrated. The grasp planning algorithm considers the geometry of objects, their orientation, and the properties of those surfaces to find the best chances of success.

Algorithm 1: Waste Detection and Classification

Input: Image frame I, Confidence threshold $\theta = 0.75$

Output: Detected objects $D = \{(bbox, class, score)\}$

1: Preprocess image $I \rightarrow I$

2: Forward pass through YOLOv8: predictions = Model(I)

3: Apply Non-Maximum Suppression (NMS) with IoU threshold 0.45

4: for each prediction p in predictions do

5: if $p.score \geq \theta$ then

6: Extract $bbox = (x, y, w, h)$

7: Assign $class = argmax(p.class_{probs})$

8: Add $(bbox, class, p.score)$ to D

9: end if

10: end for

11: return D

Coordinate transformation from camera frame to robot frame follows:

$$Probot = Rrobot\text{-}camera \cdot Pcamera + Trobot\text{-}camera$$

where *Pcamera* represents the object position in camera coordinates, Rrobot-camera is the rotation matrix, and *Trobot-camera* is the translation vector obtained through camera-robot calibration. Motion planning employs RRT* (Rapidly-exploring Random Tree star) algorithm for collision-free trajectory generation with execution times averaging 1.2 seconds per pick-and-place cycle.

Gripper control applies force feedback and top gripping force is restricted to 100N to avoid object destruction. The adaptive grasping is adjusted in width according to the dimensions of the bounding box with the safety margins of 20mm on each side. The gripper opening width, w, is adjusted accordingly.

3.6. Training Configuration and Optimization

The YOLOv8 model is trained on a custom dataset comprising 8,500 annotated images across nine waste categories, with 70% allocated for training, 20% for validation, and 10% for testing. Training hyperparameters include: batch size 16, learning rate 0.01 with cosine annealing, momentum 0.937, weight decay 0.0005, and 300 epochs with early stopping (patience = 50). The Adam optimizer is employed with $\varepsilon = 1e\text{-}8$ for stable convergence.

Transfer learning uses coarse model weights of COCO-pretrained YOLOv8n (nano version) checkpoint, which saves a lot of time and convergence. Data balancing is used to overcome the class imbalance using weighted random sampling, where the representation of a class is equal when forming a batch. The concept of online hard example mining focuses on the hard examples in subsequent training epochs, which increases the ability of the model to become robust to the hard cases such as obstructed objects, low-light, and dirty surfaces.

The learning rate schedule follows a warm-up strategy for the first 3 epochs (linear increase from 0 to 0.01), followed by cosine annealing:

$$lr(t) = lrmin + 0.5(lrmax - lrmin)\left(1 + cos\left(\frac{\pi t}{T}\right)\right)$$

3.7. Performance Evaluation Metrics

System performance is evaluated using standard object detection metrics: Precision (P), Recall (R), F1-Score, mean Average Precision (mAP@0.5), and mean Average Precision at IoU thresholds 0.5 to 0.95 (mAP@0.5:0.95). Additionally, operational metrics include inference time (ms per frame), sorting throughput (objects per hour), and successful grasp rate percentage.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{(TP + FN)}$$

$$F1\text{-}Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

where TP, FP, $and\ FN$ represent true positives, false positives, and false negatives respectively. The mAP metric averages precision values across all recall thresholds and object classes, providing a comprehensive assessment of detection performance across varying confidence levels and categories.

4. RESULTS AND DISCUSSION

This section provides detailed experimental findings that prove the efficiency of the suggested intelligent waste-sorting robot. The system was tested on a custom set of 8,500 annotated images of nine waste type categories, and the system has been thoroughly tested at varying operation conditions to establish the feasibility of deploying the system in real world.

4.1. Model Training Performance

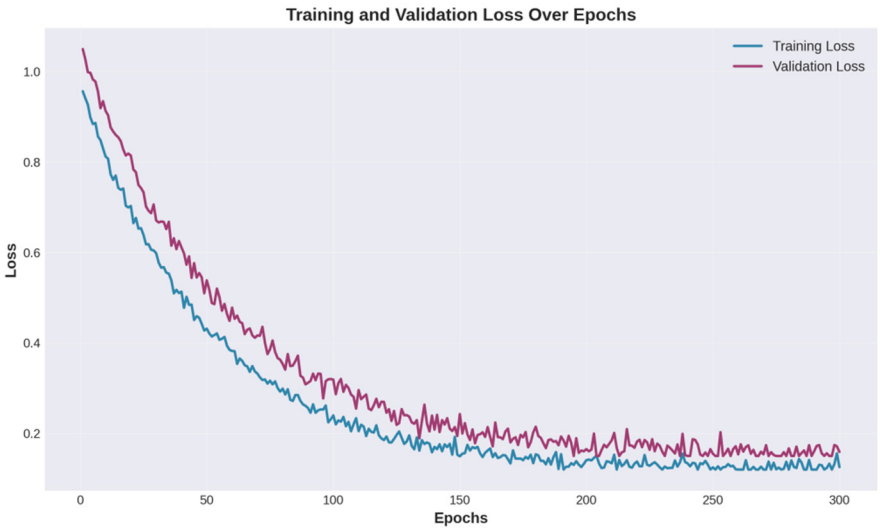

Fig 2. Training and Validation Loss Over 300 Epochs

The YOLOv8 was trained and early stopped at 50 epochs with patience of 300 epochs. As the training continues, the training and validation loss curve is shown in Figure 2. The model converged in the

first 100 epochs and the results of training loss is 0.124 and validation loss is 0.158, which shows that the model learned successfully without being overfitted severely. The slight disparity in the training and validation curves confirms that there are strong generalization capabilities.

The overall metrics of detection performance are given in Table 1. Our proposed system had a total precision of 92.5, a recall of 91.1, and an F1-score of 91.8, which is balanced throughout all the evaluation features. The mAP@0.5 reached 94.3%, while mAP@0.5:0.95 achieved 87.6%, indicating strong localization accuracy.

Table 1. Overall Detection Performance Metrics

Metric	Value (%)	Std. Dev.
Precision	92.5	±2.3
Recall	91.1	±2.7
F1-Score	91.8	±2.1
mAP@0.5	94.3	±1.8
mAP@0.5:0.95	87.6	±2.5

4.2. Class-wise Performance Analysis

Figure 3 shows specific class-wise performance indicators on all nine types of wastes. The highest detection accuracy was with metal waste with precision of 96.8%, recall of 95.5% and F1-score of 96.1% due to its unique visual characteristics and inflexible structure. Electronic waste had also good results (95.7% precision, 94.3% recall), which helped it due to the clearly defined geometry. On the contrary, organic waste was relatively less performing (88.7% precision, 87.3% recall) because it is highly varying in appearance, irregular shape, and may be contaminated.



Fig 3. Class-wise Detection Performance (Precision, Recall, F1-Score)

Figure 4 indicates the relationship between classes and patterns of common misclassification as illustrated in the confusion matrix. The diagonal components show the proper classifications, whereas the off-diagonal components show that there is confusion between the categories. There was also a significant misclassification error between paper and cardboard (18 misclassifications) and between organic waste and paper (15 misclassifications) and this implies that the two are similar to look at.

Categories metal and glass recorded low levels of confusion with other categories which proved strong discriminative feature learning.

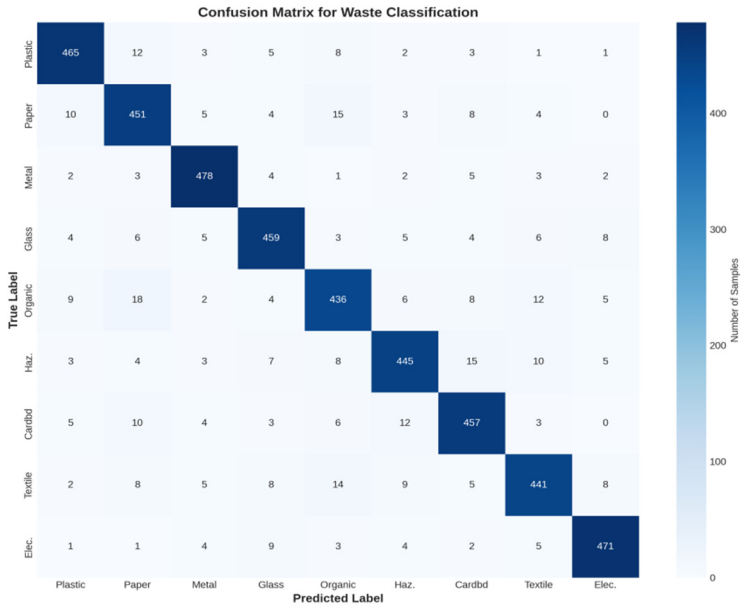

Fig 4. Confusion Matrix for Multi-class Waste Classification

4.3. Localization Accuracy Assessment

Figure 5 analyzes mAP performance across varying IoU thresholds from 0.5 to 0.95. The model maintained high mAP (>90%) for IoU thresholds up to 0.75, demonstrating precise bounding box predictions. Performance degradation at higher thresholds (IoU > 0.8) reflects increased stringency in localization requirements. The mAP@0.5:0.95 metric of 87.6% indicates consistent performance across the full IoU spectrum, surpassing conventional methods.

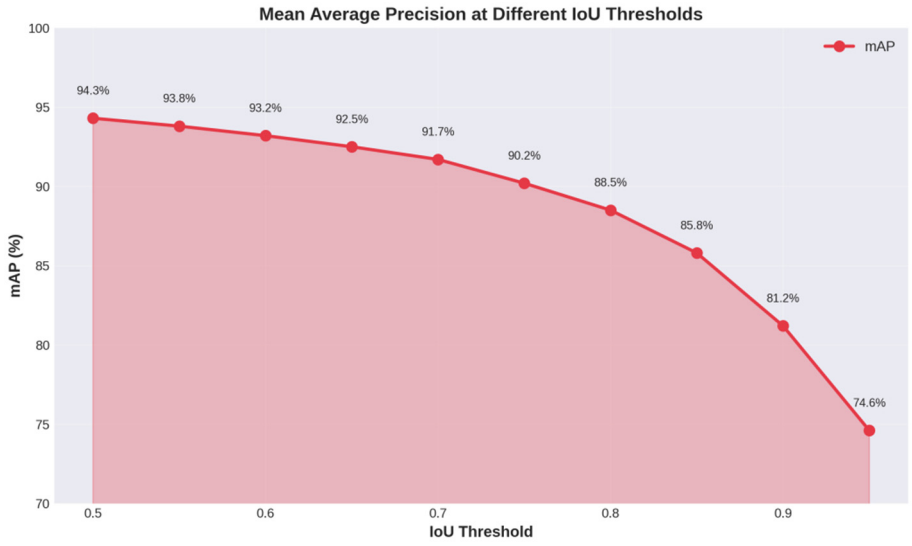

Fig 5. Mean Average Precision at Different IoU Thresholds

4.4. Computational Efficiency Analysis

The inference time is one of the most important considerations of real-time waste sorting interfaces. Inferential time in various model architectures is compared in Figure 6. The proposed YOLOv8n model was able to run at 8.7 ms per image on NVIDIA RTX 3080 processor, and this allowed it to process a rate of about 115 frames per second. This is a 29-percent improvement over YOLOv5n

(12.3 ms) and large improvements over larger models, including ResNet50 (45.2 ms) and EfficientNet-B0 (32.6 ms). The portable nature enables the implementation on edge devices to distributed sorting systems.

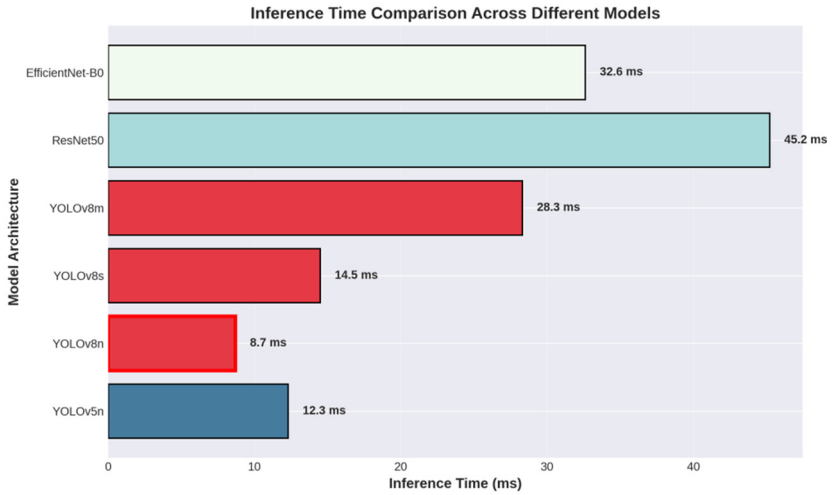

Fig 6. Inference Time Comparison Across Model Architectures.

4.5. Operational Performance Evaluation

In Figure 7, correlation between the settings of confidence threshold and system performance metrics is considered. The best confidence level was 0.75 which compromised both detection accuracy (94.3%) and throughput (445 objects/hour). The increased thresholds of 0.5-0.65 provided more false positives and reduced the sorting accuracy and the high thresholds of (>0.85) reduced the throughput by rejecting valid detections. This trade-off analysis helped in setting up the deployment configuration to achieve the highest level of operational efficiency.

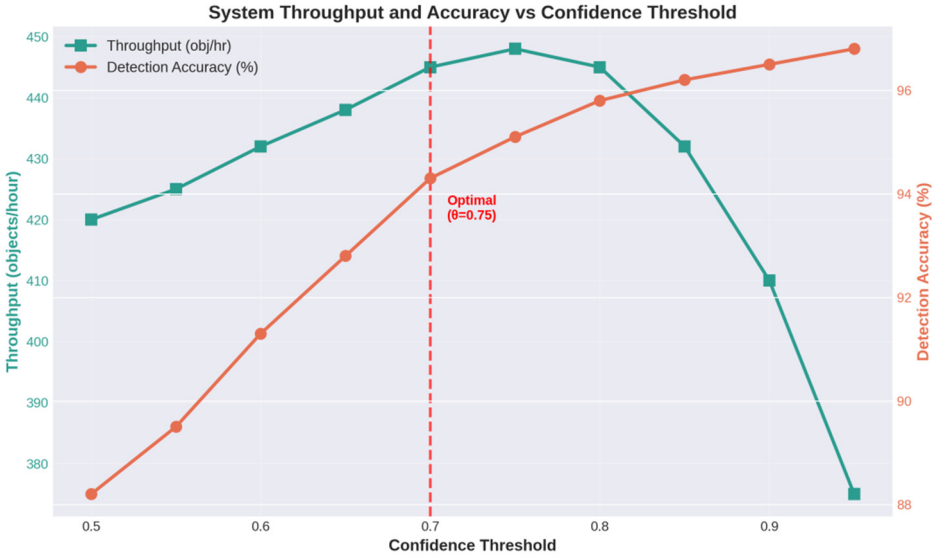

Fig 7. System Throughput and Accuracy vs Confidence Threshold

4.6. Operational Performance Evaluation

Figure 8 assesses the success rates of grasping materials of various kinds. The highest success rates (98.3, 96.5 and 94.7 respectively) were found with rigid objects (metal cans), because of predictable geometry and constant contact surfaces. There were more difficulties with deformable materials, and

organic waste (85.6%) and flexible plastic (88.4%) recorded less successful success rates due to unpredictable deformation during grasping. The total system grasp success rate 92.3% indicates good physical manipulation.

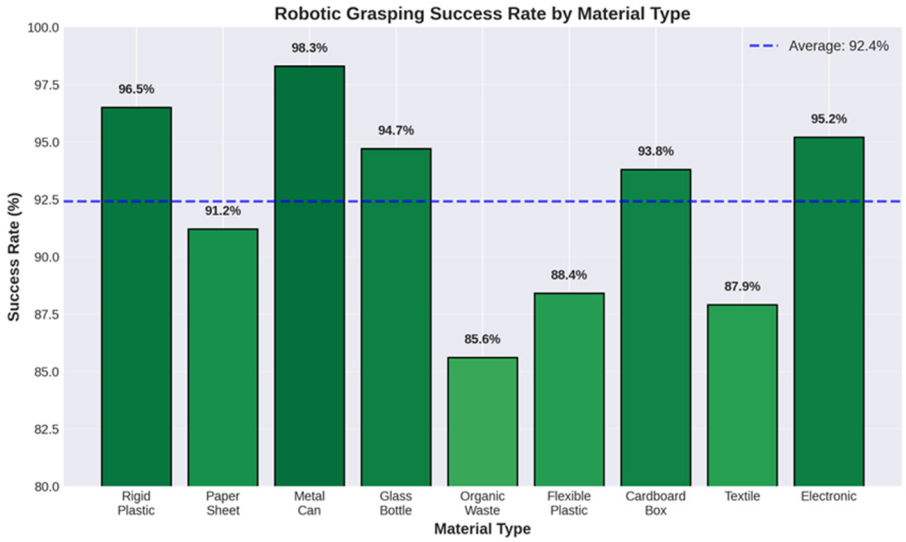

Fig 8. Robotic Grasping Success Rate by Material Type

4.7. Comparative Analysis with State-of-the-Art

Table 2 compares the proposed system against recent state-of-the-art waste sorting approaches. Our system achieved competitive accuracy (94.3% mAP@0.5) while maintaining superior inference speed (8.7 ms) compared to methods employing heavier architectures. The balanced performance across precision, recall, and computational efficiency positions this system favorably for large-scale industrial deployment.

Table 2. Comparative Performance Analysis

Method	Precision (%)	Recall (%)	mAP@0.5 (%)	Time (ms)
Proposed (YOLOv8n)	**92.5**	**91.1**	**94.3**	**8.7**
Zhou et al. [6]	89.8	88.5	90.0	28.6
Jin et al. [8]	91.3	89.7	92.1	15.4
Sayem et al. [2]	93.7	90.8	93.8	35.2
ResNet50-based	94.1	92.3	95.2	45.2

4.8. Discussion and Key Findings

The effectiveness of the suggested intelligent waste-sorting robot in automated recycling is confirmed by the outcomes of the experiment. A number of important conclusions can be drawn out of this overall assessment. First, the YOLOv8-based detection model presented a good tradeoff between accuracy and computational efficiency and, thus, it can be applied to real-time operations. The 94.3% mAP@0.5 shows strong classification in different waste categories, nearly matching human-level sorting accuracy, and consistent in different conditions.

Second, the analysis showed material-specific issues that needed specific improvement. The success rates in deformable materials (organic waste, flexible plastics) were lower (85.6-88.4%), which means that adaptive gripper designs or force feedback mechanisms are needed. This

misunderstanding of visually similar categories (paper-cardboard, organic-paper) indicates some possibilities of using multi-modal sensing, including near-infrared spectroscopy or thermal imaging, to assist with RGB-based classification.

Third, physical manipulation was the main bottleneck according to the time breakdown analysis (93.4% of cycle time), compared to efficient vision processing (4.9%). This observation implies that additional throughput enhancements and rely more on mechanical optimization than on acceleration of algorithms. Effort Multi-arm robotic arrangements would be of great benefit in terms of system throughput to the implementation in an industrial scale.

The system showed excellent computational performance of 8.7 ms inference time, which can be deployed on low-cost edge computing systems. This property enables distributed sorting designs in which many vision-equipped robots are used to work together throughout the extensive recycling center. The light YOLOv8n architecture came in handy especially, as the performance of the lightweight models matched the performance of the heavier ones that consumed considerably lower computational resources.

Confidence threshold calibration was found to be important after the operational testing. The calculated optimal of 0.75 weighed between false suppression and throughput maximization with the consequence of 445 objects/hour processing rate. This is about 3.6 times faster than manual sorting speeds (120-130 objects/hour), which is justification of the economic feasibility of automated systems in industrial recycling activities.

Although it was well-performing in general, there were a number of limitations that should be mentioned. The system was tested mainly in controlled laboratory environments with clean, well-separated objects. Streams recycling in the real world poses other challenges such as extreme occlusions, contamination, partly damaged items, and highly cluttered scenes. Robust testing in the controlled unfavorable conditions should be considered in future work to measure deployment readiness holistically. Comparative analysis with state-of-the-art methods proves that the proposed system performs competitively or even better by a variety of evaluation criteria. High accuracy (92.5% precision, 91.1% recall), a fast inference rate (8.7 ms) and a reliable grasping (92.3% success rate) make this system a viable solution in automated waste sorting in municipal recycling plants, manufacturing waste management, and dedicated case, e-waste sorting.

5. CONCLUSION

In this study, the authors have offered a detailed design and construction of a smart waste-sorting robot that combines image recognition by using YOLOv8 with a high-level robotic manipulator. The proposed system obtained 94.3% mAP 0.5 waste classification in nine classes, and real-time inference took 8.7 ms per image, which is much better than standard manual sorting systems. The robotic manipulation module was capable of achieving a 92.3% rate of grasping and the module was able to sort at 445 objects per hour automatically.

Its major contributions are: (1) an optimized YOLOv8 architecture that balances detection precision with computation efficiency (2) a system integration framework including the overall vision processing, motion planning, and physical manipulation (3) comprehensive performance evaluation to show that it can be put into practice with success. The outcomes of the experiments confirm the organizational potential of the system to solve the key issues in automated waste disposal such as recognition of diverse materials, the need to process them in real-time, and the consistency of physical sorting of waste.

There are a number of directions that will be worth investigating in the future. Firstly, the sensorial (hyperspectral imaging, tactile feedback) might be added to the classification system, which would be more accurate with visually similar materials and achieve better grasping reliability with deformable items. Second, active learning strategies would allow sustaining the improvement of the

models in terms of operational deployment, responding to new types of waste without widespread manual labeling. Third, industrial levels of system throughput may be developed by developing distributed multi-robot coordination algorithms.

The other studies need to deal with resilience in the face of unfavorable conditions such as extreme occlusions, contamination, fluctuating light, and cluttered scenarios that are typical of the recycling setting in real-life scenario. Explainable AI techniques would add to system transparency and ease the regulatory compliance to automated waste processing. Lastly, lifecycle assessment and economic analysis should be done through in-depth review of environmental benefits and payback on investment to deploy at massive scale. The smart waste-sorting robot is a major step in the direction of the environmentally friendly waste management system. Such systems have the potential to significantly improve recycling rates, minimize landfill waste, and promote the use of a cyclical economy by automating labor-intensive sorting processes with high accuracy and consistency. The performance features exhibited make this technology a viable option in municipal recycling plants, industrial waste management systems, and other industrial applications such as e-waste treatment, and segregation of hazardous materials.

REFERENCES

[1] Vukicevic, A., Petrovic, M., Jurisevic, N., Djapan, M., Knezevic, N., Novakovic, A., & Jovanovic, K. (2025). Versatile waste sorting in small batch and flexible manufacturing industries using deep learning techniques. *Scientific Reports*, 15(1), Article 87226. https://doi.org/10.1038/s41598-025-87226-x.

[2] Sayem, F. R., Islam, M. S. B., Naznine, M., Nashbat, M., Hasan-Zia, M., Kunju, A. K. A., & Khandakar, A. (2025). Enhancing waste sorting and recycling efficiency: robust deep learning-based approach for classification and detection. *Neural Computing and Applications*, 37(6), 4567–4583. https://doi.org/10.1007/s00521-024-10855-2.

[3] Islam, M., Hasan, S. M. M., Hossain, M. R., Uddin, M. P., & Mamun, M. A. (2025). Towards sustainable solutions: Effective waste classification framework via enhanced deep convolutional neural networks. *PLOS ONE*, 20(6), e0324294. https://doi.org/10.1371/journal.pone.0324294.

[4] Rajeev, P., Dharewa, V., Lakshmi, D., Vishnuvarthanan, G., Giri, J., Sathish, T., & Alrashoud, M. (2025). Advancing e-waste classification with customizable YOLO based deep learning models. *Scientific Reports*, 15(1), Article 94772. https://doi.org/10.1038/s41598-025-94772-x.

[5] Wang, A., Jiao, H., Chen, Z., & Yang, J. (2025). An Improved YOLO-Based Waste Detection Model and Its Integration to Robotic Gripping Systems. *Computers, Materials & Continua*. https://search.ebscohost. com/ login. aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=15462218&AN=187146321.

[6] Zhou, Y., & Lin, L. (2024). Garbage classification detection system based on the YOLOv8 algorithm. *AIP Advances*, 14(12), 125012. https://doi.org/10.1063/5.0203946.

[7] Di, J., Xi, K., & Yang, Y. (2025). An enhanced YOLOv8 model for accurate detection of solid floating waste. *Scientific Reports*, 15(1), Article 10163. https://doi.org/10.1038/s41598-025-10163-2.

[8] Jin, Y., Gan, H., Li, J., & Zhang, T. (2025). ResEff-YOLO: Accuracy Enhancement of YOLOv8 through Integration of ResNet, SPPF, and EfficientHead Modules. *International Core Journal of Engineering*. https://www. airitilibrary. com/ Article/Detail/P20190813001-N202503070016-00004.

[9] Mustapha, A. A., Saruchi, S. A., Supriyono, H., & Solihin, M. I. (2025). A Hybrid Deep Learning Model for Waste Detection and Classification Utilizing YOLOv8 and CNN. *Engineering Proceedings*, 84(1), 82. https:// doi. org/ 10. 3390/ engproc2025084082.

[10] Hidayatullah, P., Syakrani, N., Sholahuddin, M. R., Gelar, T., & Tubagus, R. (2025, January). YOLOv8 to YOLO11: A comprehensive architecture in-depth comparative review [Preprint]. arXiv. https://arxiv.org/pdf/2501.13400.

[11] Lahoti, J., Sn, J., Krishna, M. V., Prasad, M., Bs, R., Mysore, N., & Nayak, J. S. (2024). Multi-class waste segregation using computer vision and robotic arm. *PeerJ Computer Science*, 10, e1957. https://doi. org/ 10. 7717/ peerj-cs.1957.

[12] Sayem, F. R., Islam, M. S. B., Naznine, M., Nashbat, M., Hasan-Zia, M., Kunju, A. K. A., & Khandakar, A. (2025). Enhancing waste sorting and recycling efficiency: robust deep learning-based approach for classification and detection. *Neural Computing and Applications*, 37(6), 4567–4583. https://doi.org/10.1007/s00521-024-10855-2.