

Development of an Intelligent Waste Identification System Based on the YOLOv11 Algorithm

Sherylin Nicole Gunawan¹, Usman Sudibyo²

* Department of Informatics Engineering, Universitas Dian Nuswantoro, Semarang, Indonesia
xueryl21@gmail.com¹, usman.sudibyo@dsn.dinus.ac.id²

Article Info

Article history:

Received 2026-03-12

Revised 2026-04-22

Accepted 2026-04-30

Keyword:

*Waste Classification,
YOLOv11,
Object Detection,
Organic Waste,
Inorganic Waste,
Computer Vision.*

ABSTRACT

Effective waste management is a critical challenge in urban environments, necessitating the development of automated systems for efficient waste classification. This study aims to develop an intelligent system for detecting and classifying waste into organic and inorganic categories using the You Only Look Once (YOLO)v11 object detection algorithm. The proposed system identifies the presence and location of waste objects through bounding boxes and classifies them based on their visual characteristics to facilitate real-time detection. Experimental results demonstrate the model's effectiveness, achieving an overall accuracy of 0.717. Performance analysis indicates consistent reliability, with "Anorganik" waste achieving an accuracy of 0.708 and "Organik" waste reaching an accuracy of 0.725. Notably, despite the moderate accuracy, the model demonstrates significant robustness when deployed in real-world conditions; it is highly capable of identifying multiple objects within a single frame and maintains consistent detection performance even when objects are overlapping or clustered, confirming its viability for practical, real-time sorting applications. Despite these promising results, the study identifies several influencing factors, including object clustering, visual material similarity, lighting conditions, and camera-to-object distance. The current prototype faces limitations due to dataset size, particularly in detecting uncommon waste objects. Future development efforts will focus on expanding the training dataset to include a wider variety of waste items and enhancing robustness for detecting small or occluded objects in real-world scenarios.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. INTRODUCTION

Waste is residual material that appears after an activity or process is complete and is no longer reused [1]. The existence of waste is basically the result of human activities and natural processes that produce waste materials. Waste generally has no economic value and is considered useless in everyday activities [2]. Based on its physical form, waste can be found in various forms, such as solid, liquid, or gas, depending on its source and formation process [3]. Waste is generally divided into 2 types: organic waste and inorganic waste [4]. Inorganic waste is a type of waste that is mostly generated from human activities and has characteristics that make it difficult to decompose naturally [5].

Unlike organic waste, inorganic materials take a very long time to decompose, and in some cases, this process can take

hundreds of years [6]. These are caused by the structure of inorganic waste materials, which not easily decomposed by the microorganisms in the environment. As a result, the accumulation of inorganic waste has the potential to cause long-term environmental pollution if not managed properly [7].

Based on the study conducted by Diah Widiyarsari et al. in 2025, AdaBoost was selected as the primary model due to its superior cross-validation performance and balanced evaluation metrics, including precision, recall, and F1-score. The model's outputs were subsequently utilized to formulate a rule-based classification scheme by extracting threshold values from the most influential sensor features, thereby ensuring reliable and data-driven decision-making. Experimental evaluation demonstrated that this proposed system achieved an accuracy of 91.67% in classifying organic

and inorganic waste [8]. However, despite these strong analytical results, the practical application of the device faces several critical limitations. The system exhibits high sensitivity to environmental fluctuations, particularly lighting, and struggles to maintain accuracy when samples are in motion. Furthermore, the classification remains restricted to a narrow range of waste types, and composition detection is notably influenced by the color and material of the packaging, such as the variance between red and white plastic bags. These factors suggest that while the device is effective in a controlled, static environment, it currently lacks the robustness and versatility required for high-speed, diverse, or real-world industrial waste sorting applications.

Research conducted by Achmad Riduan et al. (2024) demonstrates that the Convolutional Neural Network (CNN) method, utilizing the ResNet101 architecture, performs effectively in classifying various types of inorganic waste, such as cardboard, glass, metal, paper, and plastic. The study reported a commendable performance, with the model achieving an accuracy of 92% on training data and 90% on test data, underscoring the efficacy of deep learning in identifying specific inorganic material categories. However, while these empirical results are strong, the study is primarily limited by a lack of transparency regarding the model's generalizability and operational diversity. Although the ResNet101 architecture shows high internal performance, the research fails to address how the system handles real-world variability—such as overlapping objects, inconsistent lighting conditions, or degraded and soiled materials—that often differ significantly from clean training datasets. Furthermore, by focusing exclusively on inorganic waste, the research does not account for the complexities of multi-material composites or organic contamination, which are inherent in practical waste streams. Finally, the absence of a discussion on computational efficiency raises questions about the model's viability for real-time deployment, particularly as ResNet101 is a deep, resource-heavy architecture that may be difficult to implement on low-power edge devices [9].

In a study conducted by Majchrowska et al. (2022), a deep learning-based framework was introduced to localize and classify litter across seven household waste categories, achieving an accuracy of up to 75% in real-world "in-the-wild" environments. By utilizing two separate neural networks—one for localization and another for identification—the proposed methodology demonstrates significant potential for waste management, offering a robust baseline for both automated mobile applications and future robotic sorting systems. However, despite these promising advancements, the framework faces several technical and practical limitations. The current system struggles significantly with the recognition of small-sized litter, indicating that while it performs well on medium and large objects, it currently lacks the precision required to surpass human-level performance. Furthermore, the authors acknowledge that industrial-scale waste sorting remains vastly more complex than household applications, requiring

the identification of more specific sub-classes, such as distinct plastic types, metal alloys, and hazardous batteries. Finally, while the model shows versatility across diverse environments—including urban, natural, and even underwater settings—the researchers note that optimizing computational efficiency and reducing predictive inference time remain critical challenges that must be addressed to ensure the system's viability for real-time deployment and large-scale environmental monitoring [10].

In this study, the researcher aims to develop an intelligent system for detecting and classifying waste into organic and inorganic categories with higher accuracy and efficiency. To achieve this objective, an artificial intelligence-based approach is applied using the You Only Look Once (YOLO)v11 object detection algorithm. The YOLOv11 algorithm is utilized to detect waste objects in images and classify them into organic or inorganic waste. Overall, the detection process in this system works as follows: first, YOLOv11 identifies the presence and location of waste objects in the image and marks them using bounding boxes. Next, the detected objects are classified based on their visual characteristics into organic or inorganic waste categories. This approach enables real-time waste detection while minimizing detection errors and improving classification performance.

II. METHOD

A. Research Framework

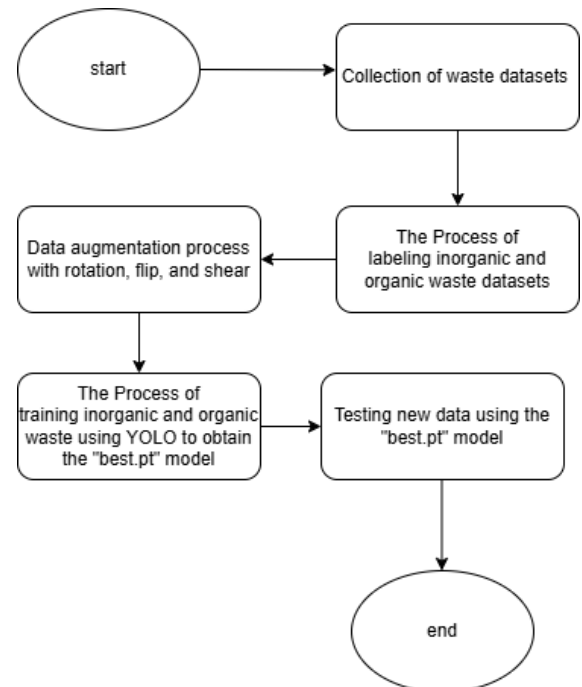


Figure 1. Research Framework

Based on Figure 1, the process begins with the initial stages:

1. **Data Collection** - The process starts with collecting image datasets containing both inorganic and organic waste. These images form the core foundation for developing the waste classification and detection system. The diversity, quality, and size of this initial dataset play a critical role in determining the final accuracy and reliability of the model.
2. **Dataset Labeling** - The collected images are carefully annotated by drawing bounding boxes around waste objects and assigning the correct class label: either "inorganic" or "organic". This labeling step transforms raw images into structured training data ready for YOLO model training.
3. **Data Augmentation** - To make the model more robust and improve its ability to generalize, the labeled dataset is expanded through augmentation techniques such as rotation, flipping, and shearing. These transformations simulate real-world variations in object orientation, viewpoint, and appearance, helping prevent overfitting and increasing overall training data variety.
4. **Model Training** - A YOLO object detection model is trained on the augmented dataset to learn to detect and classify waste as either inorganic or organic. Training is performed iteratively while monitoring performance metrics (precision, recall, mAP, loss). The model with the best validation performance is saved as best.pt.
5. **Model Testing** - The final best.pt model is evaluated using completely new, unseen test images. This phase assesses the system's real-world performance, generalization capability, and readiness for practical use in waste sorting, recycling automation, smart bins, or environmental monitoring applications.

B. Collection of Waste0 Datasets

The researchers collected the dataset independently using a smartphone camera equipped with a 48 MP Fusion main sensor, a 48 MP ultrawide sensor, and a 12 MP 5× periscope telephoto lens. The use of multiple camera sensors with different focal capabilities was intended to capture waste objects from various distances and perspectives, thereby enhancing image diversity and supporting the development of a more robust detection model. The high-resolution sensors contributed to producing images with good sharpness and detail, which facilitated the annotation process and improved the accuracy of the trained model. From the image acquisition process, the researchers successfully collected a total of 2,000 images for the waste dataset.

To improve the model's generalization ability in distinguishing between organic and inorganic waste types, the dataset was expanded using augmentation methods [11]. This process involves duplicating images using the Horizontal Flip and Vertical Flip techniques, as well as rotating them by 90°. This step aims to ensure that the system is not fixated on a

single viewing angle, but remains accurate in classifying waste even when there are changes in position or angle of the camera in a real environment [12].

The study implemented a data augmentation strategy with a ratio of 1:1.75, significantly expanding the volume of the waste dataset. By applying these transformations, the initial collection of 2,000 images was increased to a total of 3,500 images. This expansion introduces essential visual diversity, enabling the model to better generalize when identifying and classifying organic and inorganic waste under varied environmental conditions.

The final dataset of 3,500 waste images was split into training (70%), testing (10%), and validation (20%) sets, corresponding to 2,450, 350, and 700 images, respectively. This division allows for a comprehensive training process where the model's weights are optimized on the training set, while the validation and testing sets provide unbiased evaluations. Such a configuration is essential for developing a reliable classification model that performs accurately across various types of waste data.

C. Dataset Annotation

The labeling phase involves assigning specific class names to objects through bounding box annotations, a critical step for effective object detection. For this study, the Roboflow platform was utilized to manually delineate the boundaries of waste items [13]. The annotation process using Roboflow is shown in Figure 2. These annotations were subsequently exported in .txt and .yaml formats, strictly adhering to the requirements of the YOLOv11 architecture [14]. Each image in the dataset was categorized into two distinct classes: "organik" (organic) and "anorganik" (inorganic). This meticulous process ensures that during training, the model accurately associates specific visual features within the bounding boxes with their respective waste categories, providing a foundation for reliable detection as reflected in the final class distribution.

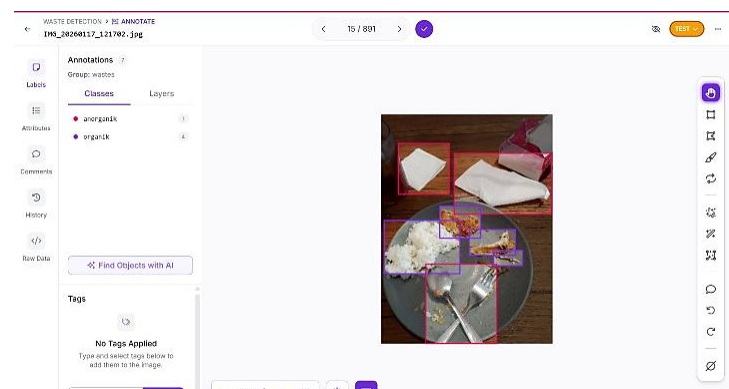


Figure 2. Annotation Process with Roboflow



Figure 3. Number of Waste Datasets

D. Data Augmentation Process

Data augmentation was implemented as a critical preprocessing step to mitigate overfitting and expand the breadth of the training samples. Although several methods—most notably shearing and enlargement—have been proposed [15], the scope of this research was focused on the application of rotation and reflection (flipping). The dataset was enriched by performing precise 90-degree rotations, both clockwise and counter-clockwise, alongside the execution of horizontal and vertical flips. This systematic variation allows the neural network to learn features regardless of their spatial orientation. Figure 4 provides a comparative illustration of the original and augmented image sets.



Figure 4. Data Augmentation Rotation and Flipping

E. Training Dataset with YOLOv11

The training phase was executed using the YOLOv11s architecture, utilizing a waste dataset partitioned into 70% training, 10% testing, and 20% validation subsets. To ensure computational efficiency and reduce processing time, the training was conducted on the Google Colab platform, leveraging the hardware acceleration of a T4 GPU. The model configuration included 30 epochs, a batch size of 16, and an input image resolution of 640x640 pixels. The YOLOv11s model employed in this study consists of 319 layers and 9,428,566 parameters, requiring 21.6 GFLOPs for its operations.

YOLOv11 hyperparameters are divided into two primary categories: training dynamics and data augmentation. Training parameters like Learning Rate (\$lr_{0}\$), Momentum, and Weight Decay govern the optimization process, determining how quickly and accurately the model converges toward a minimum loss. Additionally, "Gains" for box and classification losses help the model prioritize between localizing objects precisely and identifying their categories correctly. During inference, thresholds for Confidence and Intersection over Union (IoU) are utilized to filter out weak predictions and eliminate redundant overlapping boxes through Non-Maximum Suppression (NMS).

To improve generalization and prevent overfitting, YOLOv11 employs aggressive spatial and color augmentations. Techniques such as Mosaic, which combines four images into one, and Mixup, which blends two images, force the model to detect objects in varied contexts and scales. Other hyperparameters like HSV adjustments, Rotation, and Scaling further enhance the model's robustness by simulating different environmental conditions and camera angles. Together, these settings allow the model to remain high-performing across diverse real-world datasets without requiring manual architectural changes.

TABLE 1
PROPOSED YOLO11 ARCHITECTURE

Layer	Previous layer	n (Repetition)	Params	Module	Arguments
0	-1	1	928	ultralytics.nn.modules.conv.Conv	[3, 32, 3, 2]
1	-1	1	18560	ultralytics.nn.modules.conv.Conv	[32, 64, 3, 2]
2	-1	1	26080	ultralytics.nn.modules.block.C3k2	[64, 128, 1, False, 0.25]
3	-1	1	147712	ultralytics.nn.modules.conv.Conv	[128, 128, 3, 2]
4	-1	1	103360	ultralytics.nn.modules.block.C3k2	[128, 256, 1, False, 0.25]
5	-1	1	590336	ultralytics.nn.modules.block.Conv	[256, 256, 3, 2]
6	-1	1	346112	ultralytics.nn.modules.block.C3k2	[256, 256, 1, True]
7	-1	1	1180672	ultralytics.nn.modules.conv.Conv	[256, 512, 3, 2]
8	-1	1	1380352	ultralytics.nn.modules.block.C3k2	[512, 512, 1, True]
9	-1	1	656896	ultralytics.nn.modules.block.SPPF	[512, 512, 5]
10	-1	1	990976	ultralytics.nn.modules.block.C2PSA	[512, 512, 1]
11	-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
12	[-1, 6]	1	0	ultralytics.nn.modules.conv.Concat	[1]
13	-1	1	443776	ultralytics.nn.modules.block.C3k2	[768, 256, 1, False]
14	-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
15	[-1, 4]	1	0	ultralytics.nn.modules.conv.Concat	[1]
16	-1	1	127680	ultralytics.nn.modules.block.C3k2	[528, 128, 1, False]
17	-1	1	147712	ultralytics.nn.modules.conv.Conv	[128, 128, 3, 2]
18	[-1, 13]	1	0	ultralytics.nn.modules.conv.Concat	[1]
19	-1	1	345472	ultralytics.nn.modules.block.C3k2	[384, 256, 1, False]
20	-1	1	590336	ultralytics.nn.modules.conv.Conv	[256, 256, 3, 2]
21	[-1, 10]	1	0	ultralytics.nn.modules.conv.Concat	[1]
22	-1	1	1511424	ultralytics.nn.modules.block.C3k2	[768, 512, 1, True]
23	[16, 19, 22]	1	820182	ultralytics.nn.modules.head.Detect	[2, [128, 256, 512]]

F. YOLO (You Only Look Once) Algorithm Architecture

The You Only Look Once (YOLO) framework, originally introduced by Redmon et al. in 2015, represents a paradigm shift in real-time object detection. Since its inception, the algorithm has undergone rapid and continuous iterative development, culminating in the sophisticated YOLOv12 architecture released in 2025 [16]. As depicted in the schematic in Figure 5, the operational principle of YOLO involves partitioning the input image into an $S \times S$ grid. Within this spatial framework, the algorithm evaluates each individual grid cell to determine the presence of an object. For cells identified as containing a target, the model generates a bounding box accompanied by a confidence score (P_c), which serves as a metric for the model's certainty regarding the

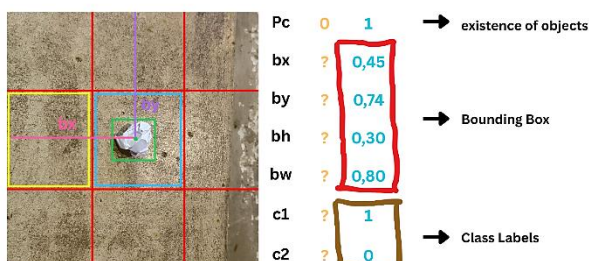


Figure 5. How YOLO works [17]

detection. Each predicted bounding box is characterized by a five-dimensional vector: $[P_c, b_x, b_y, b_h, b_w]$. In this context, P_c quantifies the likelihood that an object exists within the box, while b_x and b_y represent the precise coordinates of the center point relative to the boundaries of the specific grid cell. The dimensions b_h and b_w correspond to the normalized height and width of the bounding box, expressed as a ratio relative to the total dimensions of the image. As the YOLO lineage has evolved, the architectural complexity has significantly increased; each subsequent version introduces distinct structural modifications, such as the integration of advanced convolutional parameters, novel processing modules, and optimized residual connections to enhance feature extraction and overall accuracy [17].

As illustrated in Figure 6, which highlights the structural differences between the YOLOv11 architecture and its predecessors, this latest iteration enhances object detection efficiency and accuracy by replacing the C2f backbone

module with C3k2 and integrating C2PSA to optimize spatial focus while reducing parameter counts [18]. To address the challenges of low-contrast residues in single-crystal furnace pipelines, the architecture incorporates the C2PSA_iEMA module—combining memory-based attention (EMA) with the iRMB structure—to stabilize feature extraction against complex backgrounds. Furthermore, the BFAM_EMA module aggregates multi-scale temporal features to capture fine-grained morphological details, while the Adaptive Spatial Feature Fusion (ASFF) module is

integrated into the detection head to dynamically weigh spatial contributions across different scales. Collectively, these improvements, supported by depth-wise convolutions (DWConv) in the classification head, significantly boost the

model's robustness and precision in detecting multi-scale targets within demanding industrial environments.

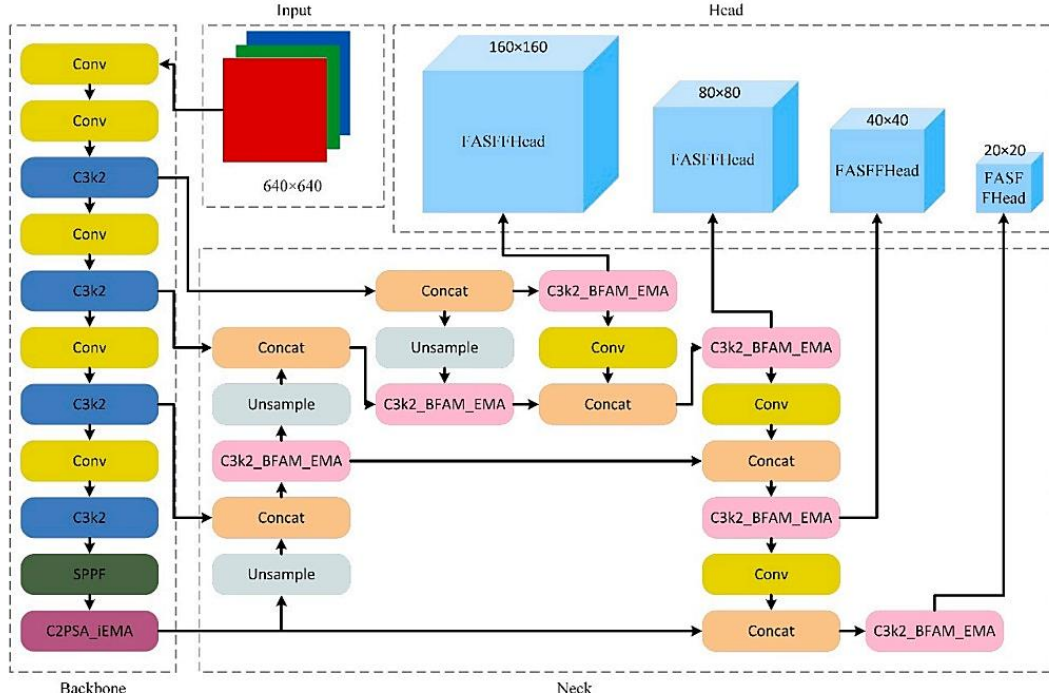


Figure 6. YOLOv11 Architecture [18]

G. Evaluation of the YOLO (You Only Look Once) Matrix

YOLO has several evaluation metrics for object detection. Here are some metrics that are often used by YOLO:

1) *Intersection of Union (IoU)*: To determine how accurately a detection box covers the ground truth, the IoU metric is applied. Based on the reference [19], the calculation for this metric is presented in Equation (1).

$$IoU = \frac{\text{The slice area}}{\text{The combined area}} \quad (1)$$

2) *Average Precision (AP)*: This metric is frequently employed in object detection tasks to evaluate model performance. It determines the area under the precision-recall curve across different threshold levels [20]. The mathematical representation for computing AP is provided in Equation (2).

$$AP = \sum_{i=1}^n (R_i - R_{\{i-1\}}) P_i \quad (2)$$

3) *Recall (R)*: To determine the sensitivity of the detection, this metric assesses how many ground truth objects were successfully captured by the model. Following the methodology in [21], the formula is defined as shown in Equation (3)

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{\text{All Ground Truth}} \quad (3)$$

4) *Precision (P)*: To assess the accuracy of positive classifications, this metric measures the extent to which the model's predictions align with reality relative to the total prediction count. The formula for this evaluation, based on [22], is shown in Equation (4).

$$Precision = \frac{TP}{TP+FP} = \frac{TP}{\text{All prediction}} \quad (4)$$

5) *Mean Average Precision at 0.5 IoU (Map50)*: Model performance is evaluated using mAP50 to measure detection success at a 0.5 IoU threshold, alongside mAP50-95 for a more rigorous assessment [23]. While mAP50 balances classification and localization, mAP50-95 provides a stricter measure of spatial accuracy by averaging precision across multiple IoU thresholds (0.5 to 0.95) [24].

$$mAP50 = \frac{1}{N} = \sum_j^K AP_{50,i} \quad (5)$$

III. RESULTS AND DISCUSSION

In this section, the researcher will present the results of the waste detection model, which classifies garbage into organic and inorganic categories. The detection was performed using

the YOLOv11 algorithm. Furthermore, an analysis of the experimental data obtained during the research will be conducted.

A. Analysis of MAP results from the waste dataset

In this study, waste detection was performed using the YOLOv11s algorithm to classify garbage into organic and inorganic categories. The model was trained for 40 epochs on a dataset containing 386 validation images with 1240 instances. The evaluation results indicate promising performance, with an overall precision (Box P) of 0.763 and a recall (Box R) of 0.699, demonstrating the model's capability to accurately identify waste objects while capturing a substantial portion of the relevant instances.

Specifically, the model showed distinct performance metrics for each class, with the anorganik class achieving a precision of 0.73, a recall of 0.677, and a mAP50 of 0.744, while the organik class achieved a precision of 0.797, a recall of 0.694, and a mAP50 of 0.72. The overall mAP50 value of 0.779 indicates strong detection performance at an IoU threshold of 0.5, and the mAP50-95 value of 0.525 reflects the model's capability across stricter IoU levels. Furthermore, with an inference speed of 5.4ms per image on a Tesla T4 GPU, the model is efficient for potential real-time applications.

TABLE 2
MAP RESULTS FROM WASTE DATASET

Class	Images	Box (P)	Box (R)	mAP50	mAP50-95	Instance
all	386	0.763	0.699	0.738	0.525	1240
Anorganik	203	0.73	0.677	0.744	0.51	396
Organik	259	0.797	0.694	0.72	0.541	844

NB :

- Box (P) : Precision
- Box (R) : Recall

B. Dataset Overview and Model Classification Performance

The model's performance was evaluated using a confusion matrix across three classes: Organic, Inorganic (Anorganik), and Background. The dataset of 3,500 augmented images was split into training, validation, and testing sets for robust evaluation.

The model demonstrated strong classification capability, achieving 666 correct predictions for organic waste and 301 for inorganic waste. Error analysis revealed that the most common misclassifications occurred between waste objects and background, with 254 background instances incorrectly classified as organic and 165 organic samples misidentified as background.

Notably, inter-class confusion between organic and inorganic waste was minimal, with only 13 organic samples misclassified as inorganic and 4 vice-versa. This indicates the model successfully learned distinguishing features between

the two waste types. The main area for improvement is enhancing the model's ability to differentiate foreground objects from complex backgrounds to reduce false detection rates.

TABLE 2
MODEL PERFORMANCE

Category	Description / Metric	Value
Dataset Distribution	Original Dataset	2,000
	After Augmented	3,500
	Train (70%)	2,450
	Test (10%)	700
	Validation (20%)	350
Classification Accuracy	True Positive (Organik)	666
	True Positive (Anorganik)	301
Error Analysis	Organik Misclassified as Anorganik	13
	Organik Misclassified as Background	165
	Anorganik Misclassified as Organik	4
	Anorganik Misclassified as Background	91
	Background Misclassified as Organik	254
	Background Misclassified as Anorganik	149

C. Analysis of bounding box prediction results

The detection results show that the model is capable of distinguishing between organic and inorganic waste fairly well. For example, a bowl of rice is correctly identified as "organik" with a 90% confidence score, while plastic bottles and hygiene products are classified as "anorganik" with confidence values ranging from 0.4 to 0.8. Each object is generally detected in proportion to its physical size, indicating that the bounding box mechanism is functioning effectively for waste management applications.

However, prediction accuracy is influenced by several factors, including object clutter from overlapping items, visual similarity between organic and inorganic materials, and varying lighting conditions or camera distance. These factors can result in overlapping detection boxes or lower confidence scores (around 0.3–0.5). Despite these limitations, the model demonstrates a robust ability to localize and classify waste in real-time settings, making it a promising solution for automated waste sorting systems.

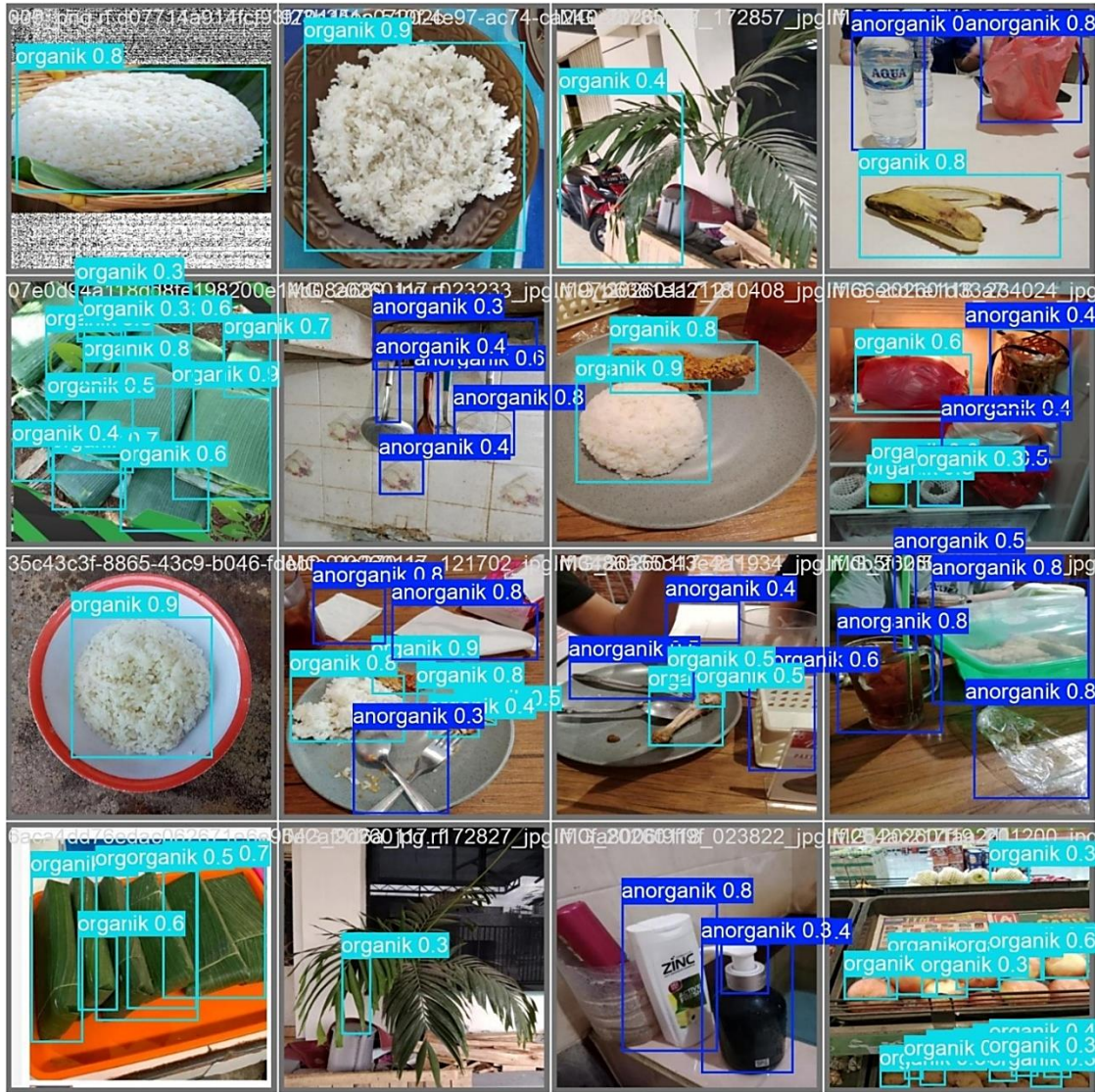




Figure 7. Predicted inorganic waste results with bounding box

TABLE 3
WASTE RECOGNITION BASED ON YOLO11

Original Image	Result Image	Output	Status
		1 Organik	True

		<p>1 Anorganik, 2 Organik</p>	<p>True</p>
		<p>2 Anorganik</p>	<p>True</p>
		<p>2 Organik, 1 Anorganik</p>	<p>True</p>
		<p>5 Organik</p>	<p>True</p>

		<p>1 Organik</p>	<p>True</p>
		<p>4 Anorganik, 3 Organik</p>	<p>True</p>
		<p>9 Organik</p>	<p>True</p>

IV. CONCLUSION

The detection results show that the model is effective in classifying waste, achieving an overall mAP50 of 0.799 and an mAP50-95 of 0.525 on our waste dataset. Specifically, the model performs reliably across both classes, with "Anorganik" waste achieving a mAP50 of 0.755 and "Organik" waste achieving a mAP50 of 0.804. These results confirm the effectiveness of the YOLOv11 algorithm in identifying and categorizing different types of waste, which is highly promising for the development of automated waste management systems.

Through this research, we identified several factors influencing detection accuracy, including object clustering, visual similarities between materials, lighting conditions, and the distance between the camera and the waste. It is important to note that our current system is a prototype with specific limitations. Due to the restricted size of our training dataset, the model struggles to detect objects that are uncommon or rarely encountered in the dataset. Furthermore, the system is currently optimized for controlled environments where waste is clearly visible. For future development, the model could be enhanced by expanding the dataset to include a wider variety of waste items and improving its robustness in detecting small or heavily occluded objects in real-world scenarios.

REFERENCES

- [1] F. Arifin, M. Habiburrahman, and W. R. Gusti, "Classification of Organic and Inorganic Waste Types Based on Neural Networks," *Elinvo (Electronics, Informatics, and Vocational Education)*, vol. 8, no. 1, pp. 78–85, Jun. 2023, doi: 10.21831/elinvo.v8i1.53284.
- [2] I. N. Chazanah, A. Bayu, and D. Nandiyanto, "Literature of Waste Management (Sorting of Organic and Inorganic Waste) Through Digital Media in Community," 2022.
- [3] H. Durak, "Comprehensive Assessment of Thermochemical Processes for Sustainable Waste Management and Resource Recovery," Jul. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/pr11072092.
- [4] I. N. Chazanah, A. Bayu, and D. Nandiyanto, "Literature of Waste Management (Sorting of Organic and Inorganic Waste) Through Digital Media in Community," 2022.
- [5] S. Syafrudin, J. M. Masjhoer, and M. Maryono, "Characterization and quantification of solid waste in rural regions," *Global Journal of Environmental Science and Management*, vol. 9, no. 2, pp. 337–352, Mar. 2023, doi: 10.22034/gjesm.2023.02.12.
- [6] I. Luhar *et al.*, "Solidification/Stabilization Technology for Radioactive Wastes Using Cement: An Appraisal," Feb. 01, 2023, *MDPI*. doi: 10.3390/ma16030954.
- [7] B. Budjav, "Evaluation of Environmental Pollution and Waste Management Strategies on the Ecosystem," *Journal of Enterprise and Business Intelligence*, pp. 223–234, Oct. 2022, doi: 10.53759/5181/jebi202202022.
- [8] D. Widiyarsari, H. Mukhtar, W. A. Cahyadi, A. Dharma, and S. Wijaya, "Design and Implementation of an Organic and Inorganic Waste Detection System Using Capacitive, Inductive, and LDR Sensors with Rule-Based Classification," vol. 7, no. 4, pp. 605–618, 2025, doi: 10.35882/ijeemi.v7i4.133.
- [9] A. Riduan, F. Panjaitan, S. Rizal, N. Huda, and S. D. Purnamasari, "Detection of Inorganic Waste Using Convolutional Neural Network Method," *Journal of Information Systems and Informatics*, vol. 6, no. 1, pp. 290–300, Mar. 2024, doi: 10.51519/journalisi.v6i1.662.
- [10] S. Majchrowska *et al.*, "Deep learning-based waste detection in natural and urban environments," *Waste Management*, vol. 138, pp. 274–284, Feb. 2022, doi: 10.1016/j.wasman.2021.12.001.
- [11] D. Hogan Itam, E. Chimeme Martin, and I. Taiwo Horsfall, "Enhanced convolutional neural network methodology for solid waste classification utilizing data augmentation techniques," *Waste Management Bulletin*, vol. 2, no. 4, pp. 184–193, Dec. 2024, doi: 10.1016/j.wmb.2024.11.002.
- [12] J. He and K. L. E. Law, "Deep Learning Models for Rotated Object Detection in Aerial Images: Survey and Performance Comparisons," *IEEE Access*, vol. 12, pp. 180436–180457, 2024, doi: 10.1109/ACCESS.2024.3509745.
- [13] A. Atalarais, K. Saputra, S. H. Syahputra, A. Idrus, and I. Taufik, "Journal of Artificial Intelligence and Engineering Applications Automatic Waste Type Detection Using YOLO for Waste Management Efficiency," 2025. [Online]. Available: <https://ioinformatic.org/>
- [14] K. Wang, J. Liu, and X. Cai, "C2PSA-Enhanced YOLOv11 Architecture: A Novel Approach for Small Target Detection in Cotton Disease Diagnosis," 2025.
- [15] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.
- [16] L. T. Ramos and A. D. Sappa, "A Decade of You Only Look Once (YOLO) for Object Detection: A Review," 2025, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2025.3630988.
- [17] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," Dec. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/make5040083.
- [18] Q. Zhao and J. Zhu, "An Improved YOLOv11 architecture with multi-scale attention and spatial fusion for fine-grained residual detection," *Results in Engineering*, vol. 27, Sep. 2025, doi: 10.1016/j.rineng.2025.107061.
- [19] X. Luo, Z. Cai, B. Shao, and Y. Wang, "Unified-IoU: For High-Quality Object Detection," Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2408.06636>
- [20] J. Tian, Q. Jin, Y. Wang, J. Yang, S. Zhang, and D. Sun, "Performance analysis of deep learning-based object detection algorithms on COCO benchmark: a comparative study," *Journal of Engineering and Applied Science*, vol. 71, no. 1, Dec. 2024, doi: 10.1186/s44147-024-00411-z.
- [21] M. Trigka and E. Dritsas, "A Comprehensive Survey of Machine Learning Techniques and Models for Object Detection," Jan. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/s25010214.
- [22] M. Owusu-Adjei, J. Ben Hayfron-Acquah, T. Frimpong, and G. Abdul-Salaam, "Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems," *PLOS Digital Health*, vol. 2, no. 11 November, Nov. 2023, doi: 10.1371/journal.pdig.0000290.
- [23] Y. Guo *et al.*, "Bi-directional YOLOv10 with average convolution for brain tumor detection in MRI," *Brain Res. Bull.*, vol. 237, Apr. 2026, doi: 10.1016/j.brainresbull.2026.111780.
- [24] R. Ardiansyah, I. G. P. Wirarama WW, and A. Y. Husodo, "Modification Of Yolov11 Nano And Small Architecture For Improved Accuracy In Motorcycle Riders Face Recognition Based On Eye," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 5, pp. 3667–3682, Oct. 2025, doi: 10.52436/1.jutif.2025.6.5.4535.