Waste patterns based on deep learning object detection YOLOv8 in tourism areas

Budi Dwi Satoto[1*], *Faroid*[1], *Aghus* Setio Bhakti[1], *Ariant* Yusuf Wibowo[1], *Royan* Dodyt Kurniawan[1], *Budi* Irmawati[2], and *Ramaditia* Dwiyansaputra[2]

[1]Information System, Engineering Faculty, University of Trunodjoyo Madura, Telang Street PO. BOX 2, Kamal Bangkalan, East Java, Indonesia
[2]Informatics Engineering, University of Mataram, Majapahit Street No.62, Gomong, Selaparang, West Nusa Tenggara. 83115, Indonesia

Abstract. The increasing population growth and demand for disposable goods and waste production complicate the sorting and processing of hazardous, inorganic, and organic waste. Meanwhile, various waste processing techniques are needed for various types of waste, including unsafe, inorganic, and organic. This research aims to address the challenges of an inefficient waste management system by utilizing deep learning technology to help better classify waste. The contribution of the research is to use lightweight deep learning to learn waste types and obtain models. The method used is YOLOv8, a lightweight object detection algorithm for classification so that it is hoped that it can help manage waste types. The advanced architecture of YOLOv8 and its integration with frameworks such as TensorFlow and PyTorch facilitate accurate and efficient waste detection. The YOLOv8 architecture is used because it can detect objects based on frames. The dataset includes styrofoam, cardboard boxes, plastic bottles, cans, and plastic wrappers. Based on the research results, the average model accuracy was 96%, with an average error value of MSE 0.0065, RMSE 0.0806, and MAE 0.0025. The training and model creation process took ten minutes. The model was tested using experimental data with an accuracy confidence level of 85-95%. This research shows that YOLOv8 can improve waste management in the area

1 Introduction

Dispersed inorganic waste can potentially harm the primary draw of tourist destinations—the surrounding natural environment. It may lessen the experience for visitors and the aesthetic appeal. Visitors may feel uneasy and unsatisfied when trash is not adequately managed or is strewn about. It may impact reviews and the reputation of tourism sites. Untidy and unclean tourist spots can become less appealing, draw fewer tourists, and harm tourism industry revenue. Good inorganic waste management is a component of sustainable tourism. It enhances the destination's image as an eco-friendly travel destination and demonstrates a dedication to environmental preservation. Effective waste management can be a helpful tool for teaching residents and tourists the value of reducing, reusing, and recycling. Effective waste management can be a valuable tool for teaching residents and tourists the value of reducing, reusing, and recycling inorganic waste. Travel destinations can receive accreditation or recognition for adhering to international waste management standards, improving their reputation and appeal worldwide. Establish a transparent sorting method to distinguish between hazardous and organic garbage and inorganic waste. It makes the process of processing and recycling easier. Inorganic waste management in tourism can enhance visitor experiences, destination reputation, and environmental quality while holistically encouraging social responsibility and legal compliance and emphasizing sustainability [1].

Waste management is a pressing worldwide issue that calls for creative solutions to maximize resources and promote sustainability. Conventional methods cannot handle the growing amount of waste and its environmental effects. However, the development of artificial intelligence (AI) technology presents a viable way around the intricacy of waste management systems. This paper thoroughly analyzes artificial intelligence's function in trash management, covering collection, sorting, recycling, and monitoring. This handbook highlights the significance of enhancing data quality, privacy precautions, cost-effectiveness, and ethical issues while outlining each application's possible advantages and difficulties. The likelihood that AI will be integrated with the Internet of Things (IoT), machine learning advancements, the significance of collaborative frameworks, and the importance of cooperative structures and policy actions are covered. In conclusion, even though AI promises to enhance waste management procedures, addressing issues like data quality, privacy problems, and financial ramifications is critical. The revolutionary potential of artificial intelligence (AI) can be fully realized to promote effective and sustainable

* Corresponding author: budids@trunojoyo.ac.id

waste management techniques with coordinated efforts and continued study [2]

Waste management results in waste dismantling, which is accomplished by landfilling and recycling. The Internet of Things (IoT) and deep learning offer quick fixes for real-time monitoring and classifying data. This paper presents the architecture of an IoT- and deep learning-based capable trash management system. The suggested model uses convolutional neural networks (CNN), a well-liked deep learning paradigm, to provide a clever method of separating trash that can be digested from waste that cannot. This plan also presents an architectural design for a smart trash can that uses a microcontroller with numerous sensors. The suggested approach monitors data using Bluetooth and the Internet of Things. While Bluetooth facilitates short-range data monitoring via Android applications, IoT allows real-time data control from any location. To evaluate the effectiveness of the created system utility-scale (SUS), sensor data estimate, waste label classification accuracy, and model were computed and analyzed. The suggested architecture's SUS score is 86%, and its classification accuracy, based on the CNN model, is 95.3125%. However, with real-time waste monitoring, this intelligent system will be adaptable to domestic activities [3].

A significant amount of picture data has been produced by the fast-growing application of artificial intelligence in waste management due to the adoption of "Zero Waste" and Industry 4.0, together with advancements in analysis techniques. Convolutional neural networks (CNN) are becoming vital for uncovering hidden patterns in visual characteristics, especially when combined with sophisticated image processing techniques. In recent years, CNNs have been used in various intelligent waste identification and recycling (IWIR) applications. The lack of generally recognized criteria and standards for the data sets and models makes it challenging to characterize the state of the art in IWIR research since CNNs are still relatively new to environmental researchers. Thus, this review's objective is to investigate the CNN methodology and how it's used in IWIR. First, some background information on CNNs is provided. Next, a description of the sophisticated CNN models and open-source datasets employed at IWIR is provided, along with an analysis of the three primary tasks: segmentation, object detection, and classification. After that, a summary is provided for the three primary uses of CNN in IWIR: the classification of solid waste, the identification of recyclable items, and the detection of waste contamination. Lastly, the difficulties and constraints of the existing applications are examined to clarify CNN's chances in this field in the future [4].

Waste and waste management are becoming increasingly crucial to wise and sustainable growth, particularly in developed and emerging nations. The waste or waste management system comprises interconnected systems that perform various intricate tasks. Deep learning (DL) has garnered attention recently as a potential substitute for traditional computational methods in the problem-solving of garbage or waste management. Because of the focus researchers have placed on this topic, a substantial amount of research has been published, especially in recent years. Numerous thorough garbage classification and detection surveys have been conducted based on the literature.

Nevertheless, no study has examined the use of DL to address trash or waste management issues across a range of domains and emphasized the datasets available for waste detection and categorization. However, the use of DL to address waste or waste management issues across domains has not yet been studied. No study has emphasized the readily available datasets for waste detection and categorization across domains. This survey helps to accomplish this goal by reviewing different models for object and image detection and classification, as well as how they are applied to waste detection and classification problems; it also provides an organized and precise analysis of waste detection and classification techniques, and it gathers more than twenty waste data sets for reference. It challenges current practices and prospects in this area to bolster this research further. As a result, scholars in this discipline will have a strong foundation in and understanding of cutting-edge deep learning models [5].

One of the world's top garbage generators is Indonesia. Every day, trash is placed in the wrong spot or not at all, and sometimes, it is not even possible to tell what kind of trash is being thrown there. Garbage often comes in two varieties: organic garbage breaks down quickly, whereas inorganic waste takes longer. This study presents five organic and inorganic trash products to test the system's ability to recognize objects at every backdrop position and the number of objects in a single image frame. Tensorflow is used in this system, and each object has a sample of 100 photos. Each sample item is assigned an identifier based on the object's name and nature. It assigns identities to these samples to train the system to identify item forms and patterns from the data objects collected. According to the research findings, the system's average percentage of accuracy in detecting objects was 90%. The objects with the highest data accuracy, tested on banana peels, leaves, cardboard, styrofoam, shattered glass, bottles, cans, and nails, were found to have 99% accuracy. Subsequently, the author attempted to fit two or more distinct things into a single frame, yielding an average percentage value of 90%. It tried to simultaneously fit two or more comparable objects into a single frame, yielding an average value of 90%. According to the test results, the grass and bottle items are samples that do not contain reading mistakes in the system. [6].

Recycling tasks that require a lot of effort can be reduced using automatic classification robots that use efficient image recognition. Nowadays, convolutional neural network (CNN) models—like DenseNet121—sessed using a well-known benchmark dataset called TrashNet, which consists of 2,527 photos with six distinct waste categories replaces more conventional image recognition methods and are the industry standard. The performance of CNN, while data augmentation techniques can increase the accuracy of CNN-driven trash classification, fully linked CNN layers' hyperparameters are never perfectly tuned. Thus,

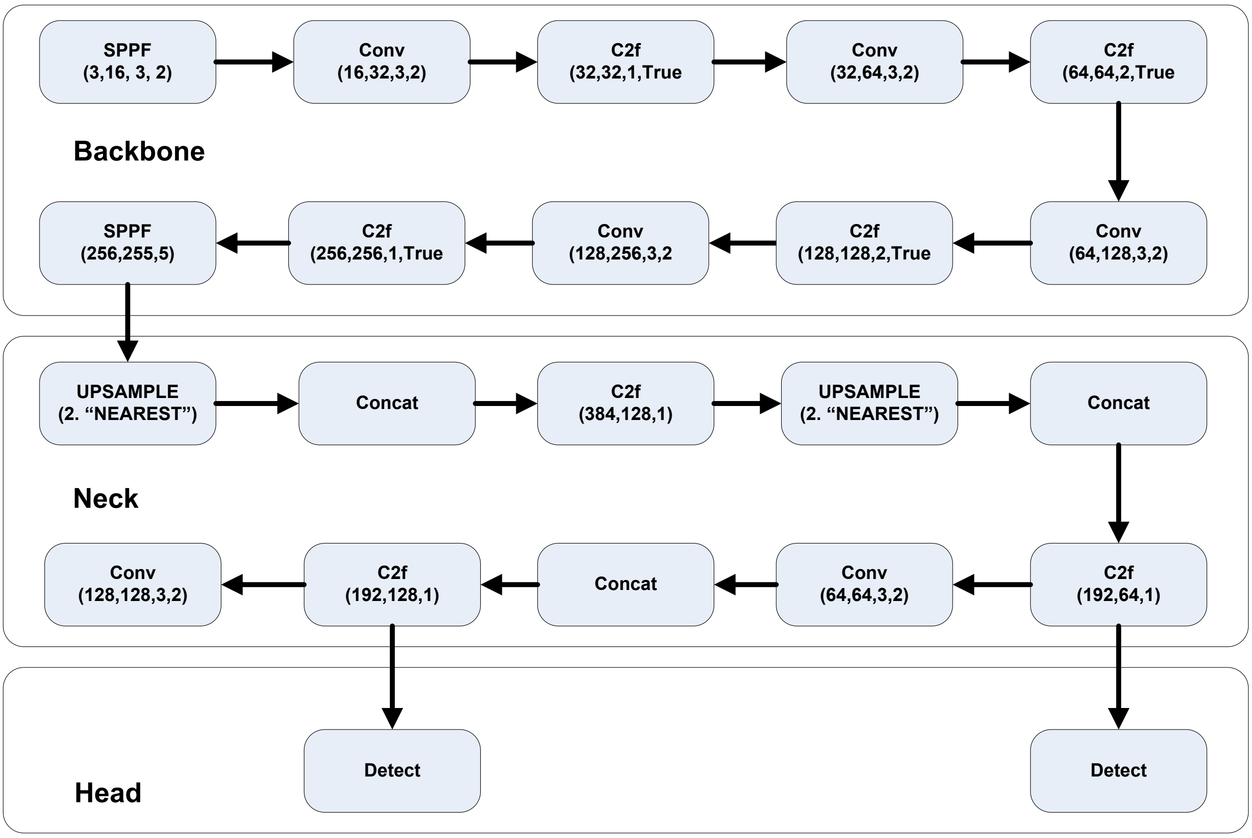


Fig. 1. Architecture of YoloV8.

in addition to data augmentation, the goal of this research is to apply a genetic algorithm (GA) to optimise DenseNet121's fully connected layers. Classification accuracy in TrashNet and provide a DenseNet121 that is optimised. According to the data, compared to other research CNNs, the optimized DenseNet121 gets the maximum accuracy of 99.6%. When it comes to increasing the accuracy of picture categorization, data augmentation can be more effective than fully connected DenseNet121 layer optimization for TrashNet. DenseNet121's performance can be improved by substituting fully connected layers for the original classifier routines. The refined DenseNet121 increases precision even more and demonstrates how effective GA is at maximizing the number of neurons and dropout rate of fully connected layers. Gradient-weighted class activation mapping adds more information on the explainability of optimised DenseNet121 and highlights coarse aspects of waste images [7].

Based on the literature presented above, the author then used a deep learning solution to select waste type patterns. This background is important to help analyze how to overcome the problem of waste types so that solutions can be found when sorting waste types. Reducing the types of waste that can be recycled from non-recyclables will help waste management. The difference with previous research is that some previous studies still use ground truth to obtain patterns, and some deep learning studies need to improve their performance to be lightweight when used. Although the identification process using ground truth is faster, the error rate is still quite significant because it requires a large number of pattern examples. A total of 400 image files per class were used in this study to obtain the best model. The research contribution is that Yolov8 has an enhanced design, which achieves greater precision and quicker inference by combining a lightweight backbone, optimised neck, and anchor-free detection head. The expected results are better precision and higher accuracy, as well as low computational costs. The urgency of this research is that this technology can provide practical solutions to support recycling systems, waste management, and environmental impact mitigation.

2 Literature review

YOLOv8 works better in utilising the head and neck regions for enhanced object recognition accuracy and more effective use of the backbone for feature extraction [8]. A multi-scale method is used to identify items in pictures of different sizes. It improves the model's ability to identify both big and small things. An improved loss function during model training helps lower detection mistakes and boost performance [9].

2.1 Deep learning object detection Yolov8

YOLOv8 incorporates enhancements to the model architecture, such as better utilization of the head and neck regions for enhanced object recognition accuracy and more effective use of the backbone for feature extraction [10]. It is a multi-scale method to identify items in images of different sizes, as shown in Fig.1. It improves the model's ability to identify small things. An improved loss function helps lower detection mistakes and boost performance [11].

$$y(i,j,k) = \sum_{m=0}^{K=1}\sum_{n=0}^{K-1}\sum_{d=0}^{D-1} x(i+m, j+n, d).w(m,n,d,k) + b(k) \quad (1)$$

Where:
$y(i,j,k)$ is the output at position (i,j) in the $k-th$ channel of the feature maps [12],
$x(i+m, j+n, d)$ is the input value at position $(i+m, j+n)$ in the $d-th$ channel [13, 14]
$w(m,n,d,k)$ is the filter value at position (m,n) in the $d-th$ channel to produce the $k-th$ channels, $b(k)$ is the bias for the $k-th$ channel. K is the filter size (e.g., for a 3x3 filter)

YOLO prioritizes efficiency and speed by completing real-time object detection in a single step. It contrasts alternative image processing methods like R-CNN, which classifies pictures using bounding boxes and CNN [15]. Bounding box prediction in Yolo:

1. Center Coordinates (x,y): The values of the center x and y coordinates fall within the range of [0, 1], as they are normalized to the grid cell size [16]. The coordinates of the grid cell's absolute center in the image are determined as follows if it is at position (i,j) in the $S \times S$ grid [17]:

$$b_x = \sigma(t_x) + c_x \quad (2)$$
$$b_y = \sigma(t_y) + c_y \quad (3)$$

Where :
t_x and t_y is the model predicted value for the center coordinates [18],
σ is a sigmoid function to ensure values are within a range [0, 1] [19],
The c_x and c_y , is the grid cell coordinate in the original image [20].

2. Dimensions of the bounding box (w,h): width (w) and height (h), normalized to the size of the picture

$$b_w = p_w e^{tw} \quad (4)$$
$$b_h = p_h e^{tw} \quad (5)$$

Where :
t_w and t_h, is the model prediction value for the bounding box dimensions [18].
p_w and p_h is the dimension of the anchor box used as a reference [21]

The likelihood that an object is inside the bounding box and the $intersection\ over\ union\ (IoU)$ between the predicted bounding box and the ground truth multiplied to yield the confidence score [22].

$$c = P(obj)xIOU_{pred,truth} \quad (6)$$

Where $P(obj)$ is the probability that there is an object in the bounding box and $IOU_{pred,truth}$ is the Intersection over Union between the predicted bounding box and the ground truth [22].

3. Object class prediction.

In addition to bounding box predictions, each grid cell also predicts the object class probability $P(classi \mid obj)$ for each possible class [23].

$$P(\,class_i \mid obj\,) = \frac{e^{t_{class_i}}}{\sum_j e^{t_{class_j}}} \quad (7)$$

Where the logit value for $class\ i$ that the model predicts is denoted by t_{class_i}. Taking an input image, processing it with CNN for feature extraction, and dividing it into grids [24]. Projecting bounding boxes and object classes in each grid cell, using Non-maximum Suppression (NMS) to remove overlaps, and finally generating results for object detection with a degree of accuracy and confidence are all part of the overall YOLO work process [25].

3 Methods

The specific improvement that YOLOv8 brings to object recognition and feature extraction is the use of Non-Maximum Suppression (NMS), as shown in Figure 2 that each object is only identified once using the most appropriate bounding box, thus reducing the number of multiple detections. The backbone extracts features ranging from simple features, such as edges, to more complex features, such as texture patterns and object shapes. YOLOv8 uses a backbone architecture that balances depth and computational efficiency. Focal Loss helps address the problem of class imbalance in the dataset. It gives more weight to the hard-to-predict and reduces the contribution from more accessible classes, thus enabling the model to focus on objects that are harder to detect.

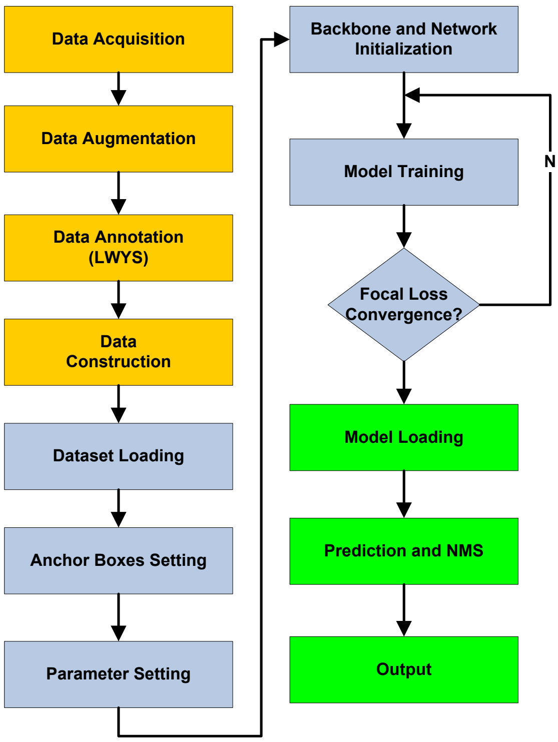


Fig. 2. Flowchart system.

3.1 Data collection

The 28,285 photos that comprised the trash data were separated into five categories: plastic wrap, cardboard boxes, cans, plastic bottles, and styrofoam. Information was gathered from the Kaggle website and travelers' observations. It shows in Fig. 3.

(a) (b) (c) (d) (e)

Fig. 3. Image data of collected waste (a) Plastic, (b) Carton box, (c) Can, (d) Bottle, (e) Styrofoam.

Creating a label on an image by adding a bounding box and class name to an item is known as annotation or object tagging. LabelImg is utilized for this operation, and the annotation format employed is the YOLO format. There are five classes for object labeling: styrofoam, cardboard boxes, cans, plastic bottles, and plastic bags. The process of labeling an object begins with the object's entire image. Following the collection of image data, the object, along with its class and data, is indicated by an object bounding box or label attached to the image. The LabelImg tool is used for labeling, and it generates data in the.txt format that contains label placement and bounding box information. This arrangement comprises format has the following elements: <width> and <height>, which are float values concerning the picture dimensions; <x_center> and <y_center>, which are the coordinates of the bounding box's center; and <object-class>, an integer identifying the class of the object.

3.2 Dataset division

The dataset of 28,285 images is divided into five classes. Each training data set (70%- 80%) is the most significant part of the dataset used to train the model. The model will learn from this data to understand the patterns in the data. Validation Data (10%-15%) is the data used to evaluate the model during the training process. The goal is to tune the hyperparameters and prevent the model from overfitting so that the model's performance remains good on new data that has never been seen. Test Data (10%-15%) is only used once the model is finished training to measure the final performance. Test data is not used in the training process or hyperparameter tuning. It can be seen in Table 1. There are also new data for a test as an experiment.

Table 1. Division of dataset.

Class	Data Training		Data Validation	
Plastic	4128	80%	1055	20%
Carton Box	4033	80%	1008	20%
Bottle	5246	80%	1311	20%
Can	4563	80%	1141	20%
Styrofoam	4640	80%	1160	20%
Total	22.610	80%	5675	20%

3.3 System requirement

Google Colab is used in the test suite as both an implementation and a teaching environment. Table 2 displays the hardware specs that were used for training:

Table 2. Google colab specification.

No	Component Needs	Specifications
1	Prosesor	Xeon 2.3GHz
2	Memory	12 s/d 52 GB
3	GPU Nvidia	A100 40GB
4	Bahasa Pemrograman	Python

4 Result and analysis

4.1 Confusion matrix

A simple visualization illustrates a classification model's performance by comparing model predictions against actual labels. A confusion matrix helps to see how well your classification model identifies different classes. Based on training results from a 28,285 dataset divided into 70% training and 30% validation data over 500 epochs. Early stopping or stopping is observed at the start of 325 epochs, as the previous 50 epochs showed no improvised performance increases. However, with an 80:20 split. It saw better performance and acceptable show in Fig. 4.

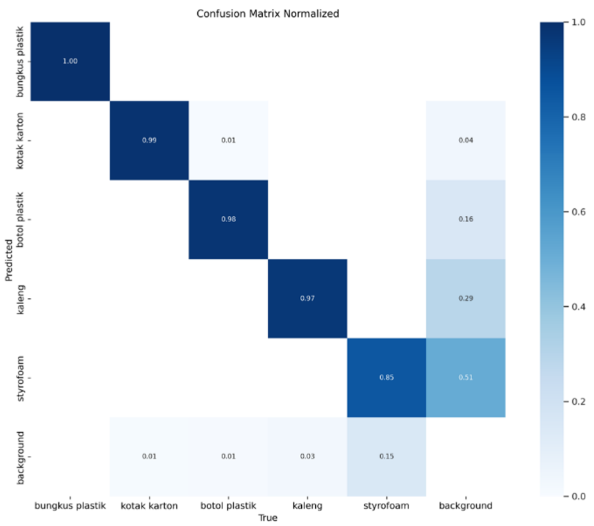


Fig. 4. Confusion matrix.

A confusion matrix will be used in the evaluation process to comprehend the performance of the trained classification model. A confusion matrix table, sometimes called an error matrix, provides an overview of how well a prediction model performs on a data collection for which the valid values are known. It allows one to see how well the model performs and recognize the many kinds of faults the model makes.

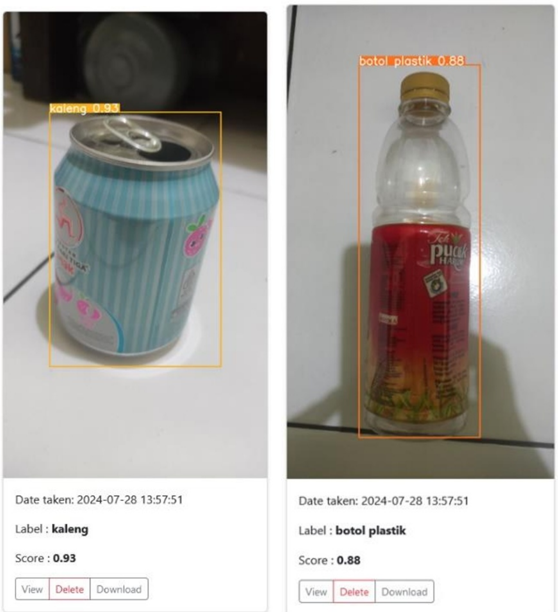
4.2 Prediction

YOLOv8 predicts bounding boxes using a technique that arranges anchor boxes on a grid. For each anchor box, the model predicts Bouncing Box Coordinates. The position (x, y), width (w), and height (h) of the bounding box.Confidence Score: The probability that the bounding box contains an object. It is calculated with a sigmoid function.

From Input, the image is received and resized. In a Backbone Network, a Convolutional network extracts visual features from the image. In the Neck Network, Features from different scales are integrated to detect large and small objects. Head Network generates bounding boxes, class predictions, and confidence scores from feature maps. Non-maximum suppression makes the Boxes overlap and filter to get the best detection. Output is Bounding boxes with class labels and confidence scores.

The output of the sigmoid function can be interpreted as a probability. For example, an output of 0.7 in binary classification can be construed as "a 70% probability that the data belongs to the positive class."

YOLOv8 combines the model's output with the anchor boxes to calculate the position and size of the bounding box. In this way, YOLOv8 can effectively detect objects in images by utilizing the information from the anchor boxes and the model predictions to generate accurate bounding boxes.



(a) (b)

Fig. 5. Tests using experimental data (a) plastic bottles, (b) cardboard boxes.

Combining objectivity probability and IoU is an effective way to evaluate the object detection performance of models like YOLOv8. This combination helps assess how well the model detects objects and accurately predicts the location and size of those objects.

According to Fig.5 test results based on experimental data, the system can accurately identify the type of garbage with a 93% confidence level for cans and an 88% confidence level for plastic bottles. It is shown in Fig.5.

The softmax function in YOLO plays a crucial role in converting raw logit scores into precise class probabilities. With softmax, YOLO can accurately classify objects in an image into possible classes, ensuring that the class prediction with the highest probability is selected. The combination of objectivity probability, class probabilities from softmax, and IoU results in a robust and efficient object detection system, which can work in real-time for applications

4.3 Precision

Accuracy is not the only factor to consider when assessing classification models. It employs two other crucial metrics: recall and precision. Recall is the percentage of genuine positives it could anticipate correctly, whereas precision is the percentage of positive optimistic predictions. At the 0.990 confidence level, the precision results for every class are 1.00 out of 5, as shown in Fig. 6. Styrofoam has the lowest precision level, and plastic wrap has the highest precision level.

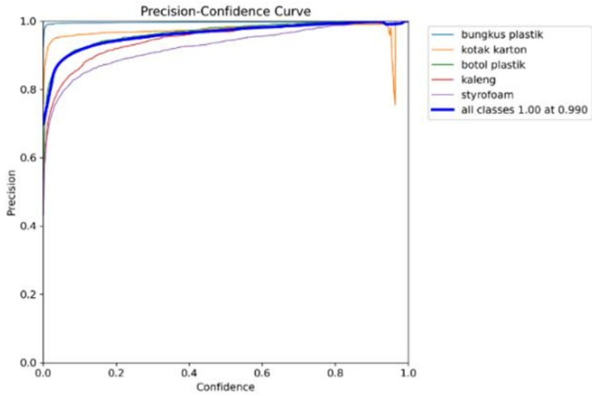


Fig. 6. Precision.

4.4 Recall

At the 0.000 confidence level, all five classes' recall results are 0.97. But there's often a trade-off between memory and precision. Enhancing accuracy may decrease recollection, and conversely. It is shown in Fig.7.

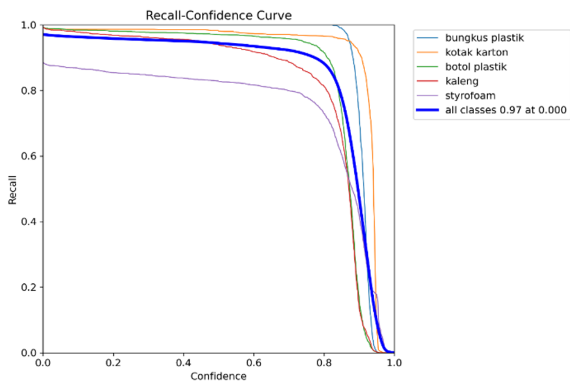


Fig. 7. Recall.

YOLO uses non-maximum suppression (NMS) to eliminate redundant or overlapping bounding box predictions. NMS helps ensure that the model only selects the best bounding box for each object, thus preventing multiple bounding boxes from detecting the same object

4.5 F-1 score

When dealing with situations where it worried about both the number of positive cases detected (recall) and the number of accurate optimistic predictions (precision), the F-1 score offers a single figure that indicates the balance between the two. An improved understanding of a model's performance on minority classes—which are frequently more complex to identify—is offered by the F-1 score shown in Fig. 8.

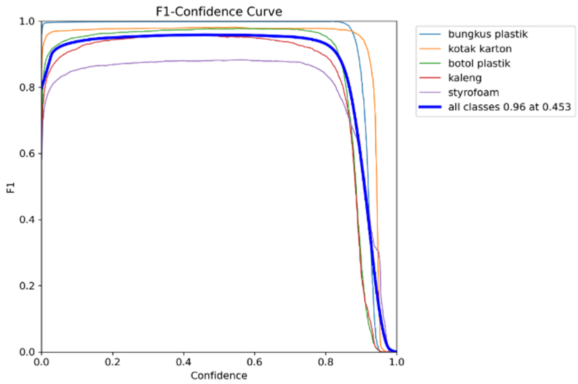


Fig. 8. F1-score.

With an average F1-score of 96%, the model was able to reach a respectable degree of accuracy through the use of this strategy.

4.6 Research comparison

Table 3. Research comparison.

No	Author	Title	Acc (%)
1	Ishita et al. (2023)	Garbage Classification Using Deep Learning [26]	95
2	Shoufeng et.al (2023)	Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling [27]	89.26
3	Sylwia et al. (2022)	Deep learning-based waste detection in natural and urban environments [28]	75
4	Abdulfattah et al. (2021)	Solid Waste Classification Using Deep Learning Techniques [29]	94.15
5	Yujin Chen et.al (2023)	Classification and recycling of recyclable garbage based on deep learning [30]	94
6	Md. Wahidur Rahman et.al (2022)	Intelligent waste management system using deep learning with IoT [3]	95
7	Current research. Budids et al. (2024)	Waste patterns based on deep learning object detection Yolov8 in tourism areas	96

5 Conclusion

The waste detection tool has been successfully used to identify various types of waste into predetermined categories, and a waste type pattern system has been successfully built. The use of the YOLOv8 method in waste identification offers several key advantages, including higher detection accuracy thanks to its anchor-free architecture and improved ability to recognize small objects, allowing for more precise differentiation of waste types. The inference speed of YOLOv8 also allows for real-time identification, making it suitable for automated sorting systems. Furthermore, this model is more efficient and easily adaptable to various environments, thus improving the performance of smart waste management applications. Based on the research results, the average model accuracy was 96%, with an average error value of MSE 0.0065, RMSE 0.0806, and MAE 0.0025. The training and model creation process took ten minutes. The model was tested using experimental data with an accuracy confidence level of 85-95%. The high level of model accuracy and the level of confidence in the test data indicate that Yolov8 deep learning and its development have great potential for use in other similar applications.

The author expresses gratitude to the National Collaborative Research program of Trunodjoyo Madura University's Community Service Research Institute

References

1. S. Wang et al., Tourists' waste reduction behavioral intentions at tourist destinations: An integrative research framework. Sustainable Production and Consumption. **25**, 540–550 (2021). https://doi.org/10.1016/j.spc.2020.12.010
2. D. B. Olawade et al., Smart waste management: A paradigm shift enabled by artificial intelligence. Waste Management Bulletin. **2**, 2, 244–263 (2024). https://doi.org/10.1016/j.wmb.2024.05.001
3. M. W. Rahman et al., Intelligent waste management system using deep learning with IoT. Journal of King Saud University – Computer and Information Sciences. **34**, 5, 2072–2087 (2022). https://doi.org/10.1016/j.jksuci.2020.08.016
4. T.-W. Wu et al., Applications of convolutional neural networks for intelligent waste identification and recycling: A review. Resources, Conservation and Recycling. **190**, 106813 (2023). https://doi.org/10.1016/j.resconrec.2022.106813
5. H. Abdu and M. H. M. Noor, A Survey on Waste Detection and Classification Using Deep Learning. IEEE Access. **10**, 128151–128165 (2022). https://doi.org/10.1109/ACCESS.2022.3226682
6. F. Fahmi and B. P. Lubis, Identification and Sorting of Waste using Artificial Intelligence Based on Convolutional Neural Network. ELTICOM. (2022). https://doi.org/10.1109/ELTICOM57747.2022.10038044
7. W.-L. Mao et al., Recycling waste classification using optimized convolutional neural network.

Resources, Conservation and Recycling. 164, 105132 (2021). https://doi.org/10.1016/j.resconrec.2020.105132

8. J. Wang et al., Road defect detection based on improved YOLOv8s model. Scientific Reports. **14**, 1, 16758 (2024). https://doi.org/10.1038/s41598-024-67953-3
9. L. Wei and Y. Tong, Enhanced-YOLOv8: A new small target detection model. Digital Signal Processing. **153**, 104611 (2024). https://doi.org/10.1016/j.dsp.2024.104611
10. F. Feng et al., Improved YOLOv8 algorithms for small object detection in aerial imagery. Journal of King Saud University – Computer and Information Sciences. **36**, 6, 102113 (2024). https://doi.org/10.1016/j.jksuci.2024.102113
11. J. Cho and K. Kim, Detection of moving objects in multi-complex environments using selective attention networks (SANet). Automation in Construction. **155**, 105066 (2023). https://doi.org/10.1016/j.autcon.2023.105066
12. R. Varghese and M. S., YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. ADICS. (2024). https://doi.org/10.1109/ADICS58448.2024.10533619
13. W. Chai et al., FDW-YOLOv8: A Lightweight Unmanned Aerial Vehicle Small Target Detection Algorithm Based on Enhanced YOLOv8. iWRF&AT. (2024). https://doi.org/10.1109/iWRFAT61200.2024.10594370
14. S. Sun et al., Multi-YOLOv8: An infrared moving small object detection model based on YOLOv8 for air vehicle. Neurocomputing. **588**, 127685 (2024). https://doi.org/10.1016/j.neucom.2024.127685
15. Z. J. Khow et al., Improved YOLOv8 Model for a Comprehensive Approach to Object Detection and Distance Estimation. IEEE Access. **12**, 63754–63767 (2024). https://doi.org/10.1109/ACCESS.2024.3396224
16. X. Wang et al., Keypoint regression strategy and angle loss based YOLO for object detection. Scientific Reports. **13**, 1, 20117 (2023). https://doi.org/10.1038/s41598-023-47398-w
17. G. Huang, B. Si, and F. Tang, Model learning based on grid cell representations. ROBIO. (2017). https://doi.org/10.1109/ROBIO.2017.8324553
18. B. Zhou et al., Object Detection and Mapping with Bounding Box Constraints. MFI. (2021). https://doi.org/10.1109/MFI52462.2021.9591174
19. B. Dave et al., Wild Animal Detection using YOLOv8. Procedia Computer Science. **230**, 100–111 (2023). https://doi.org/10.1016/j.procs.2023.12.065
20. Z. Jrondi, A. Moussaid, and M. Y. Hadi, Exploring End-to-End object detection with transformers versus YOLOv8 for enhanced citrus fruit detection within trees. Systems and Soft Computing. **6**, 200103 (2024). https://doi.org/10.1016/j.sasc.2024.200103
21. P. Vasanthi and L. Mohan, Efficient YOLOv8 algorithm for extreme small-scale object detection. Digital Signal Processing. **154**, 104682 (2024). https://doi.org/10.1016/j.dsp.2024.104682
22. L. Wang et al., Bounding Box Vectorization for Oriented Object Detection With Tanimoto Coefficient Regression. IEEE Transactions on Multimedia. **26**, 5181–5193 (2024). https://doi.org/10.1109/TMM.2023.3330103
23. M. Talib, A. Al-Noori, and J. Suad, YOLOv8-CAB: Improved YOLOv8 for Real-time object detection. Karbala International Journal of Modern Science. (2024). https://doi.org/10.33640/2405-609X.3339
24. Q. Wu et al., Bounding Box Projection for Regression Uncertainty in Oriented Object Detection. IEEE Access. **9**, 58768–58779 (2021). https://doi.org/10.1109/ACCESS.2021.3072402
25. I. Shukhratov et al., Optical detection of plastic waste through computer vision. Intelligent Systems with Applications. **22**, 200341 (2024). https://doi.org/10.1016/j.iswa.2024.200341
26. I. Joshi, P. Dev, and G. Geetha, Garbage Classification Using Deep Learning. ICCPCT. – (2023). https://doi.org/10.1109/ICCPCT58313.2023.10245133
27. S. Jin et al., Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling. Waste Management. **162**, 123–130 (2023). https://doi.org/10.1016/j.wasman.2023.02.014
28. S. Majchrowska et al., Deep learning-based waste detection in natural and urban environments. Waste Management. **138**, 274–284 (2022). https://doi.org/10.1016/j.wasman.2021.12.001
29. A. E. B. Alawi et al., Solid Waste Classification Using Deep Learning Techniques. ICOTEN. (2021). https://doi.org/10.1109/ICOTEN52080.2021.9493430
30. Y. Chen et al., Classification and recycling of recyclable garbage based on deep learning. Journal of Cleaner Production. **414**, 137558 (2023). https://doi.org/10.1016/j.jclepro.2023.137558