



JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY

<https://e-journal.uum.edu.my/index.php/jict>

How to cite this article:

Dewi, C., Susanto R. A., Chernovita, H. P., Santoso, M. V. V. & Chen, A. P. S. (2025). Evaluation of garbage detection efficacy utilising deep learning. *Journal of Information and Communication Technology*, 24(4), 36-57. <https://doi.org/10.32890/jict2025.24.4.2>

Evaluation of Garbage Detection Efficacy Utilising Deep Learning

¹Christine Dewi, ²Rafael Andi Susanto, ³Hanna Prillysca Chernovita,

⁴Melati Viaeritas Vitrieco Santoso, & ⁵Abbott Po Shun Chen

^{1,2&4}Department of Information Technology,

Satya Wacana Christian University, Indonesia

³Department of Information Systems,

Satya Wacana Christian University, Indonesia

⁵Department of Marketing and Logistics Management,

Chaoyang University of Technology, Taiwan

*¹christine.dewi@uksw.edu

²672021035@student.uksw.edu

³hanna.chernovita@uksw.edu

⁴672021106@student.uksw.edu

⁵chprosen@gm.cyut.edu.tw

*Corresponding author

Received: 10/4/2025

Revised: 3/8/2025

Accepted: 6/8/2025

Published: 31/10/2025

ABSTRACT

Garbage management has become an urgent global challenge due to the expected 70% increase in volume between 2025 and 2050, driven by rapid urbanisation and population growth. Low public understanding of garbage sorting and source-based disposal is reflected in Indonesia's inefficient garbage processing. In addition to being ineffective, traditional methods like landfill disposal and incineration present serious environmental hazards. These traditional methods are often ineffective due to resource constraints and time-consuming nature, and cannot be relied upon for extended periods because they damage the environment during the process. To increase the effectiveness and precision of garbage management, this research proposes a garbage detection and classification model using YOLOv11. This research includes data collection and pre-processing, model training, and performance evaluation using metrics such as mean average precision (mAP). This research uses the Trashnet Garbage Classification Dataset, which has 2,524 total images that are divided into six categories. The

key technical contributions of this research are to apply additional techniques, such as data augmentation strategies, to the dataset, enabling a comparison between the original and more advanced datasets. The purpose of the data augmentation technique is to improve model generalisation. The results of the evaluation metrics show that the model using the augmented dataset has slightly better performance with an mAP50 value of 97.8% than the model using the original dataset. This model is capable of identifying and classifying accurately all of the categories in the dataset.

Keywords: Computer vision, deep learning, garbage classification, object detection, YOLOv11.

INTRODUCTION

Garbage is one of the most persistent global problems right now. Around 2.12 billion tons of garbage were produced per year all over the world (Kuang et al., 2024). Of course, this phenomenon occurred due to rapid economic growth and improvements in people's living standards. A recent report from International Lianhe Zaobao shows that the volume of garbage all over the world is expected to increase by 70% between 2025 and 2050 (Kang et al., 2020; Adedeji & Wang, 2019). It is becoming a challenge itself because this is not directly proportional to the awareness of the surrounding community. The level of public awareness in Indonesia remains relatively low regarding garbage sorting, despite its importance.

Garbage is a byproduct generated by human activities on an industrial, household, and organisational scale (Ismail et al., 2023). Garbage has become a common problem in every region in Indonesia, requiring a collaboration between the government and the communities to fix and find solutions (Fitriana et al., 2024). Garbage management is a significant challenge for our country because poor garbage management contributes to the increasing accumulation of garbage (Aryantiningasih et al., 2024). Generally, garbage has two treatments, through the sorting process and without the sorting process. However, the majority of Indonesia's population still adopts the treatment without a sorting process (Ismail et al., 2023). Garbage sorting is very important to do before the garbage enters the landfills (Putri et al., 2022). Garbage sorting can be started from the smallest group, like households. The surrounding community still lacks garbage sorting and tends to dispose of garbage improperly.

Conventional garbage treatment methods, such as burning and burying, can cause significant soil and air pollution. (Chen & Xiong, 2020). The conventional methods are still applied in the community, usually starting from burning household garbage. Landfills also applied this system of burying garbage into the ground because it's considered to be. However, burying garbage has a high potential to cause soil pollution. Garbage collection and sorting are still done manually using human labour in some landfills. Surely, this thing will take a long time to separate organic garbage and inorganic garbage (Arvio et al., 2024). Classifying garbage into organic and inorganic categories is a very important part of garbage management. Proper classification allows for more efficient recycling, reduces negative impacts to the environment, and forms the basis for sustainable garbage management policies (Bahagia & Akbar, 2024).

Artificial intelligence, deep learning, and machine learning have emerged significantly to improve knowledge in various fields (Arip et al., 2024; Cao & Yongke Sun, 2024). Deep learning, a part of artificial intelligence, has become a subject of research and focus for many researchers due to its advantages in data processing. Numerous great algorithms can be applied in various industries (H. Liu, 2024). Deep learning has been applied in various fields, including education, which has led to

significant changes and unprecedented opportunities in learning, teaching, evaluation, and management (Cao & Yongke Sun, 2024). Deep learning also provides an evolution of computer vision, with a major shift from traditional methods to modern methods that rely on algorithms (Emmanuel et al., 2024).

Object detection is one of the rapidly growing fields of deep learning right now, as it allows a machine to interpret, understand, and identify objects in images from their characteristics (Khanam & Hussain, 2024). Object detection is used to identify relevant objects in a dataset and determine their category and location by using a bounding box (Fan & Song, 2024). Before the modern methods using deep learning existed, object detection used traditional methods such as Histogram of Oriented Gradients (HOG) (D. Zhang, 2024). Researchers worldwide are continuously developing various methods for garbage detection and classification, supported by computer vision technology (Zhou et al., 2024). Object detection with modern approaches is divided into two parts: one-stage detection, such as You Only Look Once (YOLO) and Single Shot Detector (SSD), and two-stage detection, such as Region-based Convolutional Neural Network (R-CNN), Fast CNN, and Faster R-CNN. Both of these use approaches based on Convolutional Neural Network (CNN) to learn existing datasets with accurate results (D. Zhang, 2024). One-stage detection uses a grid of objects to make a prediction (Carion et al., 2020). Convolutional neural networks became the first stage of object detection development with modern approaches after going through conventional approaches (Ahmed et al., 2024). These launched algorithms can improve the accuracy and speed of the detection process (Fan & Song, 2024).

One of them is the YOLO algorithm that has been developed by Redmon et al. (2015). This discovery resulted in a major contribution to the field of object detection (Khanam & Hussain, 2024). This YOLO algorithm is a very effective tool for identifying and classifying objects because it produces a high level of accuracy and efficiency of processing (Banduka et al., 2024). YOLO keeps continuing to evolve from YOLOv1 in 2015 to the latest version, namely YOLOv11. YOLOv11 was just released in September 2024. YOLOv11 has better features than previous versions by providing optimisation in efficiency and speed in processing, and higher accuracy with fewer parameters (Khanam & Hussain, 2024).

Nowadays, we need to use leading technology in this garbage classification process, along with the complex challenges in dealing with garbage problems, especially the household waste in Indonesia. The use of deep learning for object detection based on the YOLOv11 algorithm is one of the best solutions that can be utilised. With faster processing times and higher mean average precision mAP scores, YOLOv11 is a great option for real-time applications like garbage classification. These benefits overcome the drawbacks of previous iterations, such as YOLOv5 and YOLOv8, especially when it comes to detecting small or partially obscured objects. In order to improve detection performance in complex environments, YOLOv11 was chosen as the study's core model (Khanam et al., 2025). This algorithm not only produces high accuracy in the detection and classification process but also provides efficiency in terms of the computing time required in the process of garbage sorting. Surely, this technology can help to increase public awareness about the importance of garbage sorting and reduce errors in the sorting process, which is usually done manually.

The increasing need for an effective, real-time method to detect and classify various forms of household waste, particularly in difficult-to-reach places like small, obscured, or irregularly shaped objects, is what spurred this study. By utilising the most recent YOLOv11 model for multi-class garbage detection in both training and testing, this study makes a technical contribution. It compares its performance to that of YOLOv8 to show gains in speed and accuracy. According to the results, YOLOv11 is suitable for real-time applications since it achieves a higher mean average precision (mAP)50 than YOLOv8. The

efficacy of the model and its potential for incorporation into self-governing waste management systems are confirmed by these outcomes.

RELATED WORKS

There have been many previous studies related to object detection using various methods. Research by Aradea et al. (2024) about organic waste detection using Attention Convolutional Neural Network with a total accuracy of 0.8693. This research also compares the results between Attention CNN and CNN itself; the results show that Attention CNN has a slightly higher accuracy (Aradea et al., 2024). Research conducted by Kang et al. (2020) discusses an automatic garbage classification system based on deep learning using the ResNet-34 algorithm. The purpose of this research is to improve efficiency in garbage management in response to the continued global increase in waste volume. ResNet-34 algorithm obtained an accuracy of 99.96% (Kang et al., 2020).

Flores (2023) researched testing the object detection model with garbage as the dataset using the YOLOv5 algorithm. The purpose of this research is to evaluate the performance of YOLOv5 and compare it to other algorithms. The dataset used by this research includes six categories of garbage. The result of this research shows a good performance. Overall, the model has 91.6% precision, 90.2% recall, 95.2% mAP50, and 77.8% mAP50-95. However, this research did not explore data augmentation techniques or compare the results with recent models (Flores, 2023). Development of an application to detect inorganic waste using YOLOv8, as discussed by Arvio et al. (2024). The purpose of this research is to replace the traditional garbage sorting process with a modern and efficient method. Sorting manually takes a long time and is inefficient. The dataset used in this research includes five categories of inorganic garbage with a total of 2,000 images. The YOLOv8 algorithm demonstrates exceptionally high performance in detecting garbage, achieving an accuracy of 99.63%. This result shows that YOLOv8 is effective for detection and also increases the efficiency of garbage sorting. This research lacks the Arvio et al. (2024).

Lastly, the research conducted by Bahagia and Akbar (2024) discusses the development of a classification system that can support the recycling and garbage management process. The algorithm used in this research is a convolutional neural network with a dataset of garbage, which is divided into only two categories/classes, such as organic and inorganic garbage. CNN used a two-stage detection, including feature learning and classification. In the feature learning stage, convolutional and pooling will be performed. The model will have four convolutional layers, while the classification stage will use two dense layers. The parameters used were 10 epochs, 32 batch sizes, and a 0.001 learning rate. This model provides 95% accuracy in the classification process. CNN has the potential to classify objects effectively, and this process is applicable (Bahagia & Akbar, 2024).

This study develops a system for detecting and classifying garbage using the YOLOv8 algorithm, which is integrated into the autonomous vehicle to enhance automatic waste management in the urban environment. Using about 15,000 labelled images from 44 different types of household waste and data augmentation techniques, the model is refined and optimised to achieve detection accuracy of about 90% and speed up to 35 frames per second on the NVIDIA Jetson Xavier NX. The study's findings indicate that YOLOv8 can outperform previous models like YOLOv5 and SSD, particularly in detecting small or partially occluded objects in various conditions. Future research is made possible by several limitations, such as the need for more extensive testing in harsh environments, improving the diversity of datasets for better generalizability, optimising the model for deployment on devices with limited

resources, enhancing the detection of highly occluded waste, integrating the system with more extensive automated waste management processes, and carrying out long-term evaluations to guarantee operational stability in practical applications. Lack of generalisations is the gap in this research (Zhou et al., 2024).

Unlike earlier studies, this study examines the recent model YOLOv11 for multi-class garbage detection, encompassing six categories: trash, cardboard, glass, metal, paper, and plastic. This study investigates the architectural enhancements and detection capabilities of YOLOv11, in contrast to previous research that concentrated on binary classification or older YOLO versions. In order to create a more thorough baseline and assess the advantages of data augmentation, we also provide a performance comparison with YOLOv8 using the augmented datasets.

YOLOv11

As the name implies, the YOLO algorithm can find objects and their locations by processing the entire image with only one step (Khanam & Hussain, 2024). YOLO is a highly developed one-stage object detection algorithm that offers several advantages, enabling high-level accuracy in real-time object detection (Ali & Zhang, 2024). YOLO has been developed since 2015 and still exists today. YOLO has released the latest breakthrough, namely YOLOv11. YOLOv11 with significant advances for object detection. This version improves its architecture and training methodology, increasing accuracy, speed, and efficiency (Khanam & Hussain, 2024). YOLO algorithm uses a Deep Convolutional Neural Network that is able to predict the bounding box and class of the object. Computation time is significantly reduced due to more effective calculation, which is very important for real-time applications. The prediction results are based on a grid system, where each cell makes an output with a bounding box form and calculates the probability value (Singh et al., 2024).

YOLO has undergone various evolutions since its first development, from YOLOv1 to YOLOv11, with each version bringing impactful improvements in accuracy, computation time, efficiency, and inference speed. YOLOv1 introduced the grid approach to object detection, which was then improved upon by YOLOv2 and further enhanced by YOLOv3 and YOLOv4, through techniques such as an optimised backbone architecture, batch normalisation, and safety boxes. With the use of strategies like Generalised Efficient Layer Aggregation Network (GELAN), Programmable Gradient Information (PGI), and the elimination of Non-Maximum Suppression (NMS), YOLOv6 to YOLOv10 further enhanced efficiency, allowing the user to compute quicker and get more precise detection for real-time applications (Alif, 2024; Sapkota et al., 2025). YOLOv11 is the latest version of the YOLO algorithm, signalling a rapid progress in the field of object detection. Speed, efficiency, and overall performance of YOLOv11 are improved with a more efficient training pipeline and optimised design. YOLOv11 has an advanced backbone and neck architecture designed to optimise a feature. YOLOv11 offers a balanced approach, combining computational time, precision, and efficiency, making it highly suitable for a wide range of applications, from integrated systems to large-scale deployments (Sapkota et al., 2025).

Figure 1

YOLO's Evolution Process

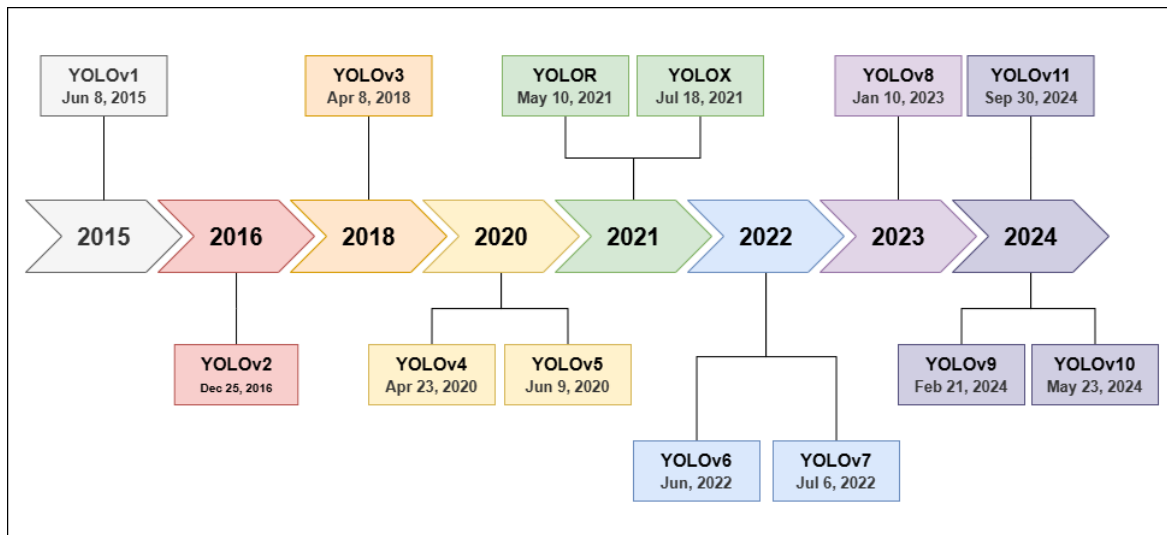


Figure 1 shows a timeline of YOLO from the first release to YOLOv11 (Jegham et al., 2024). Each version released has its capabilities and abilities, ranging from the ability to detect more datasets, the ability to increase resolution during system operation, and an architecture that is continuously updated to achieve better performance in predicting and identifying small objects. When YOLOv6 onwards was released, developers focused on improving the scalability of the model, improving the performance of real-time metrics, and reducing the computational requirements (Sapkota et al., 2024).

Compared to YOLOv8, YOLOv11 has many advantages, especially in terms of detection accuracy and inference efficiency. With an inference speed of roughly 10 ms per image and a refined architecture consisting of a CSPDarknet53 backbone, decoupled C2F head, and anchor-free design, YOLOv8 is capable of detecting small, tilted, and occluded objects. However, YOLOv11 outperforms it with a faster inference time of roughly 7.7 ms per image. It indicates that YOLOv11 is better suited for large-scale real-time applications since it not only maintains strong detection performance across a variety of object types, including rare defects, but also offers an ideal balance between accuracy and speed (Khanam et al., 2025; Zhou et al., 2024).

Figure 2

YOLOv11's Architecture

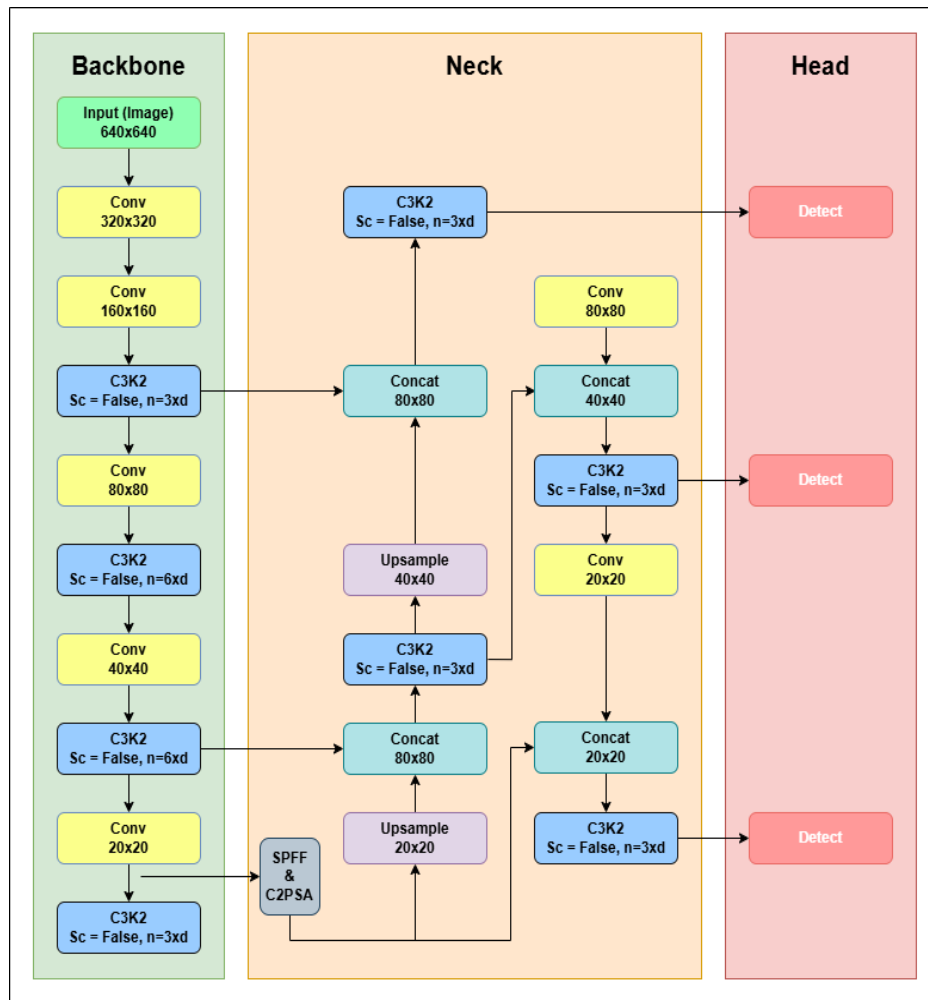


Figure 2 shows the YOLOv11 architecture that specifically optimises the feature extraction capability, which allows the model to identify both complicated and small details in the images. The goal of the YOLOv11 backbone architecture is to increase detection accuracy and improve efficiency. This model incorporates innovations such as C3K2, a lightweight cross-stage partial block that preserves feature richness while reducing computational complexity, SPPF (Spatial Pyramid Pooling Fast), which enables the model to extract multi-scale spatial features effectively; and C2PSA, a convolutional block that is integrated with Parallel Spatial Attention to highlight significant areas in the input image, all of which contribute to the improvement of the graphical fit (Vrsalovic et al., 2025). In addition, YOLOv11 is designed to achieve a balance between computational and performance capabilities, enabling efficient detection of small and large objects without compromising speed.

With this capability, YOLOv11 can identify objects with a high degree of accuracy without being affected by the orientation, scale, and size of the object. This tool makes YOLOv11 particularly useful in various industries for analysing images or videos quickly and accurately (Sapkota et al., 2024).

METHODOLOGY

Trashnet Garbage Classification Dataset

We conducted our experiment with trashnet garbage as the dataset. The dataset was taken from the GitHub repository “Trashnet Garbage Classification Dataset” (M. Yang & Thung, 2016). This dataset is divided into six categories, such as paper, glass, plastic, metal, cardboard, and trash, which represent typical household waste types. These categories were selected to represent the general waste commonly found in daily situations. The sample of the dataset will be shown in Figure 3.

Figure 3

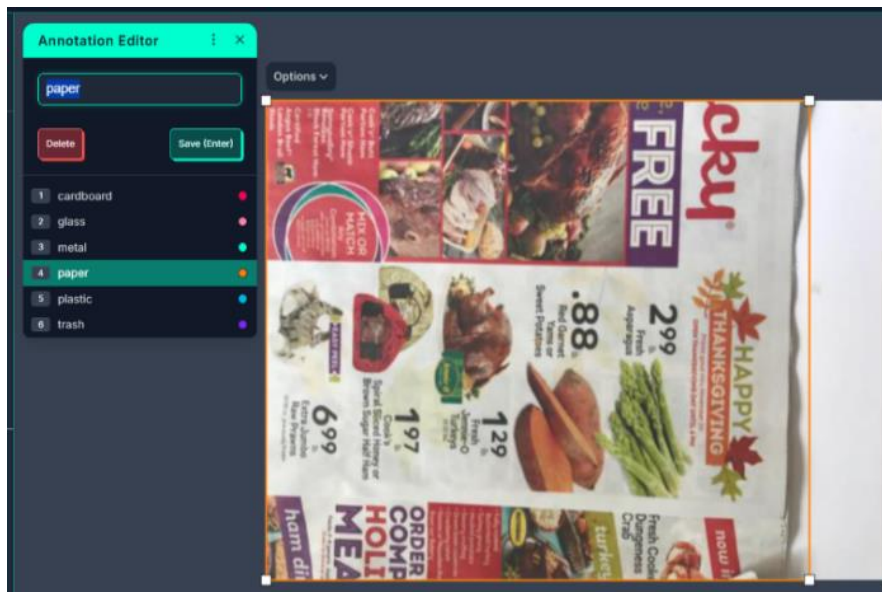
Sample Images in the Dataset: (a) Paper, (b) Glass, (c) Plastic, (d) Metal, (e) Cardboard, (f) Trash



These images will be uploaded, annotated, and pre-processed in Roboflow. Each image was manually annotated by making bounding boxes around each item. Roboflow will automatically generate annotation files in the YOLOv11 format, which includes the class index and normalised bounding box coordinates. This dataset consists of 2,524 total images with a total of 6 categories. Around 450-500 images per category, except for the trash category, which has only 104 images. A suitable dataset can optimise and support the models developed with high accuracy for various conditions and objects in the dataset.

Figure 4

Example of Annotation Process in Roboflow



The next stage is dataset pre-processing. In this research, the dataset is annotated manually to make the annotation process more accurate and consistent. This step is done with the help of Roboflow. Each image is carefully examined and annotated to ensure accurate classification. After the object has been identified, a bounding box is added around the object. The model can understand the categories better when the shadows of the object are not included in the bounding box. This process is applied to all of the existing images in the dataset. The division ratio for the dataset is 70:20:10. 70% for the training set, 20% for the validation set, and 10% for the test set. Also, there are two processes in this step such as auto-orient and resize.

Experimental Setting

We divide the dataset into two categories, Dataset 1 and Dataset 2. Dataset 1 contains the original dataset from the TrashNet garbage classification dataset without any additional techniques, with a total of 2,524 images. This dataset is used as a starting point for assessing and evaluating the model's performance without augmentation or any modifications. Dataset 2 contains a dataset that has been through data augmentation techniques to increase the number of images and enhance the model's ability to generalise the images. The purpose of this step is to make the model more resilient to changes in lighting, viewing angles, and environmental circumstances. The data augmentation techniques, such as rotation, flipping, and cropping, are used in this step. To counter any potential bias in the original dataset, Dataset 2's image count is also raised by balancing the distribution of the categories. The total number of images in Dataset 2 is 6,058.

The data augmentation process on Dataset 2 enables us to assess whether augmenting the dataset results in a quantifiable performance improvement, this justifies its inclusion in the evaluation, even though the original images in Dataset 1 were appropriate for initial training. Models were trained using a learning rate of 0.01. With the optimiser set to auto, the training framework could automatically choose the best optimisation algorithm depending on the task configuration and environment. All of the images

were resized to 640 x 640 pixels to guarantee consistency in input dimensions. To ensure that models could efficiently learn from the dataset, the training process was spread across 120 epochs.

The speed and effectiveness of model training and evaluation are greatly increased by using Google Colab with an NVIDIA Tesla T4 GPU for all experiments. This configuration, which is a cloud-based GPU resource, allows for faster computation time, optimised parallel processing, and smooth deep learning task execution. Using the high-performance architecture of the Tesla T4 allows for deeper model exploration, which not only cuts down on training time but also helps maintain high accuracy.

Figure 5

Research's Workflow

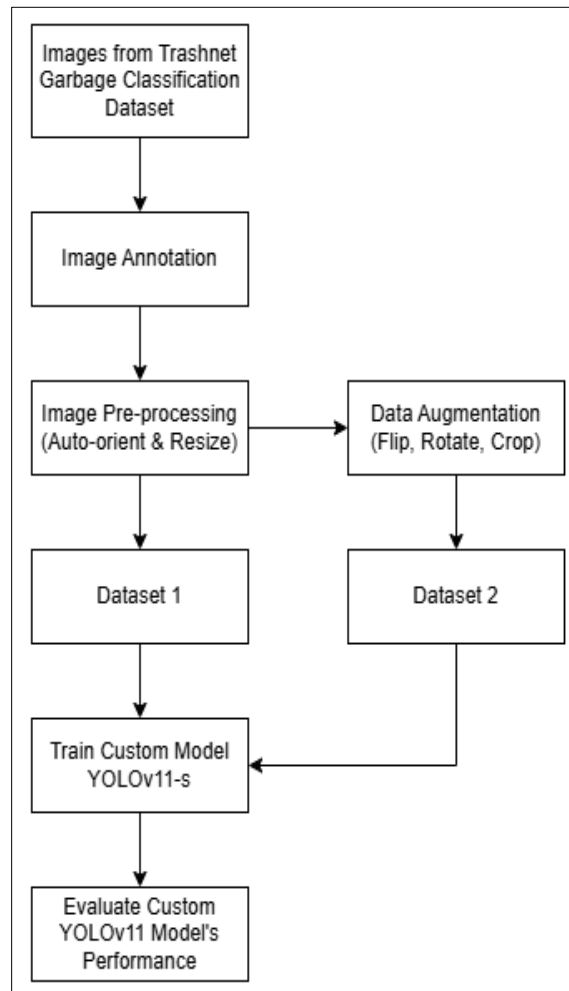


Figure 5 shows the workflow of this research. Our experiments will use YOLOv11-s to train Datasets 1 and 2. YOLOv11-s was used in this research because of its efficiency, speed, and accuracy. This model has fewer parameters, which makes it very suitable for real-time detection. This model is still quite good at identifying objects accurately.

Training Result

Initially, 60 epochs were used for the training; the results of mAP values were relatively low, indicating less than ideal performance. The training process was extended to 120 epochs to enhance the accuracy and consistency of the models. To examine and evaluate the model, important performance metrics like Precision (Equation 1), Recall (Equation 2), and mAP were monitored during the training.

$$Precision = \frac{(TP)}{(TP + FP)} \quad (1)$$

$$Recall = \frac{(TP)}{(TP + FN)} \quad (2)$$

where,

TP = True Positive,
 FP = False Positive,
 FN = False Negative.

Precision is used to measure the correct proportion of predictions made by the model, while recall is used to measure the proportion of correctly detected objects out of the total sample. Higher recall value means that the model has a high level of sensitivity (M. Zhang et al., 2025). Intersection of Unit (IoU) is used to calculate the ratio between the prediction bounding box and the actual bounding box (H. Yang et al., 2019). Equation 3 shows the calculation of IoU. Average precision is used to evaluate the overall performance of the model. mAP is the average value taken from the average precision (M. Zhang et al., 2025). Equation 4 shows the calculation of average precision, and Equation 5 shows the calculation of mAP. It is a common evaluation metric in object detection tasks. The mean average precision when the IoU threshold is set at 0.50 is known as mAP50. Accordingly, if a predicted bounding box overlaps the ground truth by at least 50%, it is considered accurate. The average mAP calculated over a range of IoU thresholds, from 0.50 to 0.95, with a 0.05 step, is indicated by mAP50–95. mAP50-95 provides a more thorough evaluation of model performance across different levels of localisation strictness, offering a more inclusive metric. In this research, we will focus on the mAP50 value as our evaluation process.

$$IoU = \frac{Area_{pred} \cap Area_{gt}}{Area_{pred} \cup Area_{gt}} \quad (3)$$

$$Average\ Precision = \int_0^1 P(r)dr \quad (4)$$

$$Mean\ Average\ Precision = \frac{1}{c} \sum_{i=1}^c AP \quad (5)$$

where,

$Area_{pred}$ = area of predicted bounding box,
 $Area_{gt}$ = area of ground truth bounding box,
 AP = average precision,
 $P(r)dr$ = area under precision-recall curve,
 c = number of object classes.

RESULTS AND DISCUSSION

Table 1 shows the result from training the YOLOv11-s model using both datasets. The metrics shown are precision, recall, and mAP50. In this research, mAP50 was used to evaluate the model's performance and determine how well it could detect and classify objects. The other metrics, such as precision and recall, are only for insight and comparison. Overall, both models performed excellently, with the model trained on Dataset 1 achieving total precision of 0.935, total recall of 0.911, and mAP50 of 0.973. A slightly lower precision of 0.922 but a higher recall of 0.961 and a total mAP50 of 0.978 were achieved by the model trained on Dataset 2 with the augmentation process, indicating that the data augmentation process enhanced the model's capacity and ability to detect more objects, even though at a slight precision trade-off. Overall, Dataset 1 has 97.3% of mAP50, while Dataset 2 has around 97.8% of mAP50. The difference between these two is slightly lower in terms of the number of images, with Dataset 2 being slightly better. When evaluated per category, the impacts of the data augmentation process are more obvious. Recall improved from 0.92 to 1 in the cardboard category, while mAP50 slightly increased from 0.981 to 0.983.

The glass category shows a noticeable improvement, with the value of recall increasing from 0.878 to 0.996 and precision decreasing from 0.996 to 0.891. It indicates that the model improved the ability to identify glass objects, albeit with a significant decrease in accuracy. With slightly greater mAP50 values, the recall for the metal and paper categories also improved to 1, indicating complete object coverage in those categories. All metrics showed improvements for the plastic category, with the precision rising from 0.962 to 0.982. However, in the trash category, recall decreased from 0.889 to 0.822, the mAP50 value decreased from 0.93 to 0.914, and the precision value increased slightly from 0.914 to 0.937. These results indicate that the data augmentation process in this category may have introduced variation that made the detection process more challenging. With all factors considered, this comparative analysis shows that the data augmentation process generally increases recall and mAP50 values, enabling the model to accurately detect more instances, particularly for specific categories like glass, metal, and paper. Although the difference in results from Dataset 1 and Dataset 2 is relatively low, the augmented dataset introduces additional variations that simulate real-world conditions. It highlights the importance of carefully considering the data augmentation process to ensure balanced improvements across all object categories.

Table 1

Training Result in Trashnet Garbage Dataset with YOLOv11-s

Class	Original (Dataset 1)			Data Augmentation (Dataset 2)		
	P	R	mAP50	P	R	mAP50
all	0.935	0.911	0.973	0.922	0.961	0.978
cardboard	0.928	0.92	0.981	0.918	1	0.983
glass	0.996	0.878	0.987	0.891	0.996	0.992
metal	0.905	0.978	0.992	0.915	1	0.994
paper	0.906	0.898	0.969	0.889	1	0.995
plastic	0.962	0.9	0.983	0.982	0.948	0.992
trash	0.914	0.889	0.93	0.937	0.822	0.914

Figure 6 shows various graphs that represent the results during the training and validation process for Dataset 2. The graphs show constant improvements over 120 epochs. The model's performance during the training process is displayed in the first row of the graphs, and validation results are displayed in the second row of the graphs. Both the training and validation loss graphs display a constant downward trend, resulting in good generalisation and showing the model is learning efficiently without showing any indications of overfitting or underfitting. The model performs well in terms of accuracy and sensitivity when detecting objects, as evidenced by the precision and recall graphs, which initially show some fluctuations before both metrics generally improve constantly.

Furthermore, the mAP50 and mAP50-95 graphs show consistent or even rising values over epochs, highlighting the model's excellent performance in precisely detecting and classifying objects as well as creating precise bounding boxes. Effective learning and good generalisation without overfitting are indicated by the loss graphs' consistent fall. Additionally, mAP50 keeps increasing, which means that the model's accuracy in detecting and identifying the object is good.

Figure 6

Training Process Dataset 2 (Data Augmentation) using YOLOv11-s

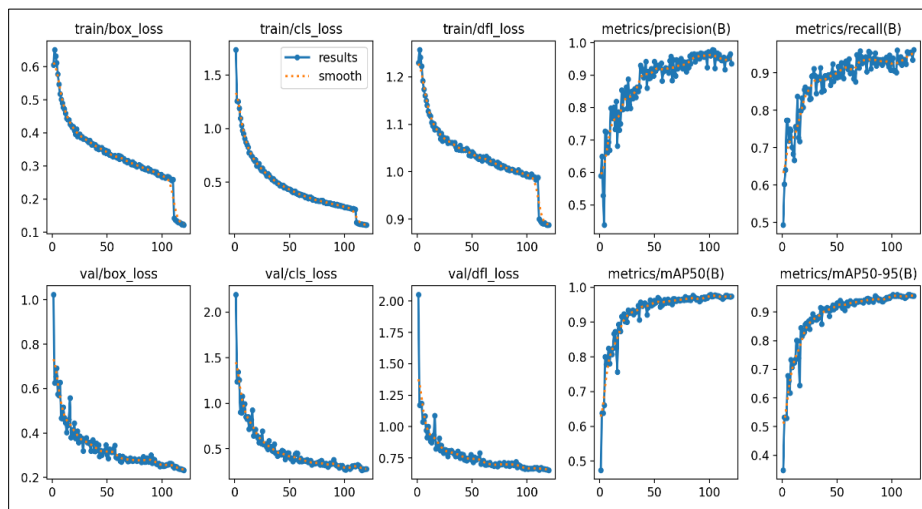


Figure 7

Validation Dataset 2 (Data Augmentation) Batch 0 using YOLOv11-s

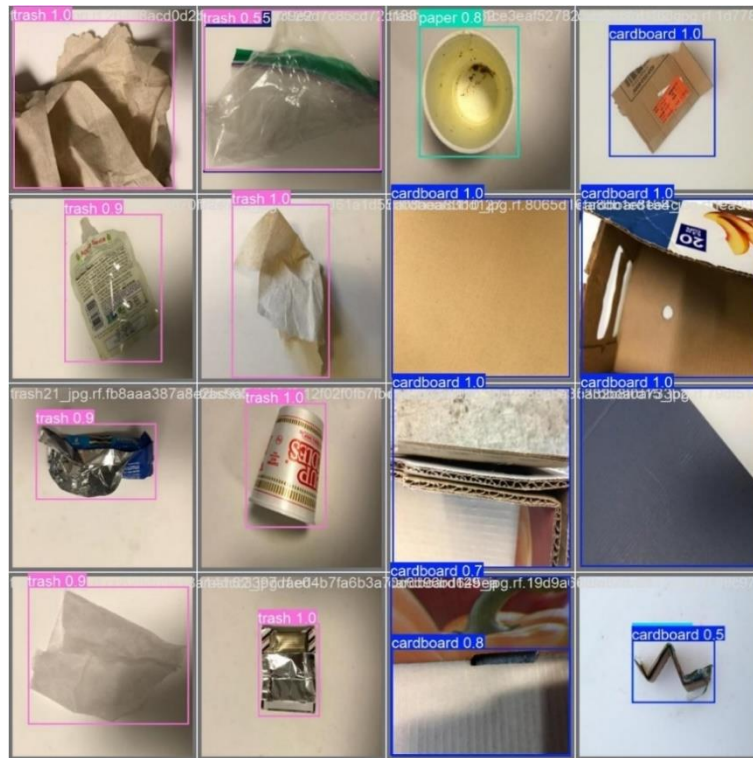


Figure 7 shows the sample of validation process on Dataset 2 where classes like trash, paper, and cardboard are successfully classified and detected. Each object is also accompanied by a confidence score that indicates how certain the model is of its classification. These outcomes show how well the model can differentiate between various waste materials, assisting in the creation of an intelligent waste sorting system by means of automated detection and classification.

Table 2

Testing Result with Trashnet Garbage Dataset using YOLOv11-s

Class	Original (Dataset 1)			Data Augmentation (Dataset 2)		
	P	R	mAP50	P	R	mAP50
all	0.935	0.911	0.974	0.921	0.961	0.978
cardboard	0.928	0.92	0.981	0.917	1	0.983
glass	0.996	0.878	0.988	0.891	0.997	0.992
metal	0.905	0.98	0.992	0.913	1	0.994
paper	0.904	0.898	0.969	0.888	1	0.995
plastic	0.962	0.901	0.983	0.982	0.948	0.992
trash	0.914	0.889	0.93	0.937	0.822	0.914

Table 2 shows the results of the testing that was conducted on both datasets. This table provides a reference for evaluating the performance of the models we have trained and tested. Overall, Dataset 2 has better performance in terms of mAP50 and recall value, with 97.8% mAP50 and 96.1% recall value.

These results indicate an excellent performance for detection and classification accuracy, with a high recall value that minimises the miss of objects. The mAP50 value from Dataset 2 is slightly higher than that from Dataset 1, with a difference of only about 0.4%. It indicates that the model trained with both datasets performed very well. The data augmentation process, which helps the model in better generalising across different object appearances, could be responsible for this minor improvement. Despite the minor improvement, this highlights how the data augmentation process can be used to improve the model’s performance without significantly changing the dataset.

Table 3

Training and Testing Result with Trashnet Garbage Dataset (Dataset 2) using YOLOv8-n

Class	Training Result			Testing Result		
	P	R	mAP50	P	R	mAP50
all	0.943	0.937	0.965	0.943	0.937	0.965
cardboard	0.934	0.9	0.976	0.935	0.9	0.976
glass	0.933	0.976	0.988	0.933	0.976	0.988
metal	0.971	1	0.994	0.971	1	0.994
paper	0.907	0.994	0.959	0.907	0.99	0.96
plastic	0.981	0.964	0.991	0.981	0.964	0.991
trash	0.934	0.791	0.883	0.934	0.79	0.884

The findings, which are displayed in Tables 2 and 3, demonstrate that Dataset 2 raised the mAP50 from 0.965 to 0.978 and enhanced the overall model’s recall from 0.937 to 0.961. The performance gains in several classes, particularly glass, paper, and metal, despite a minor decrease in precision, from 0.943 to 0.921, imply that data augmentation improves the model’s capacity to generalise by subjecting it to a wider range of circumstances. These findings support the use of YOLOv11, which demonstrated more consistent and better overall performance than the baseline YOLOv8, and validate the significance of augmentation in creating robust models.

Moreover, Figure 8 shows a sample of detection results between models using Dataset 1 and Dataset 2. Figure 8(a) illustrates that the model trained on Dataset 1 misclassified the object (paper) as trash. In Figure 8(b), the model trained on Dataset 2 detects and classifies the object correctly as paper. Some instances of models trained with dataset 1 being misclassified still exist, despite the model already performing well in many cases. It was true when an object’s visual features were not clear or the training dataset lacked sufficient variation. This indicates that a model’s performance is influenced by the variety and quantity of datasets generated through the data augmentation process. Data augmentation is important to improve the model’s generalisation, minimise overfitting, and also increase the variety of data. Furthermore, the augmentation process makes training more accurate and efficient.

Figure 8

Sample of Detection Result using YOLOv11-s for Both Datasets

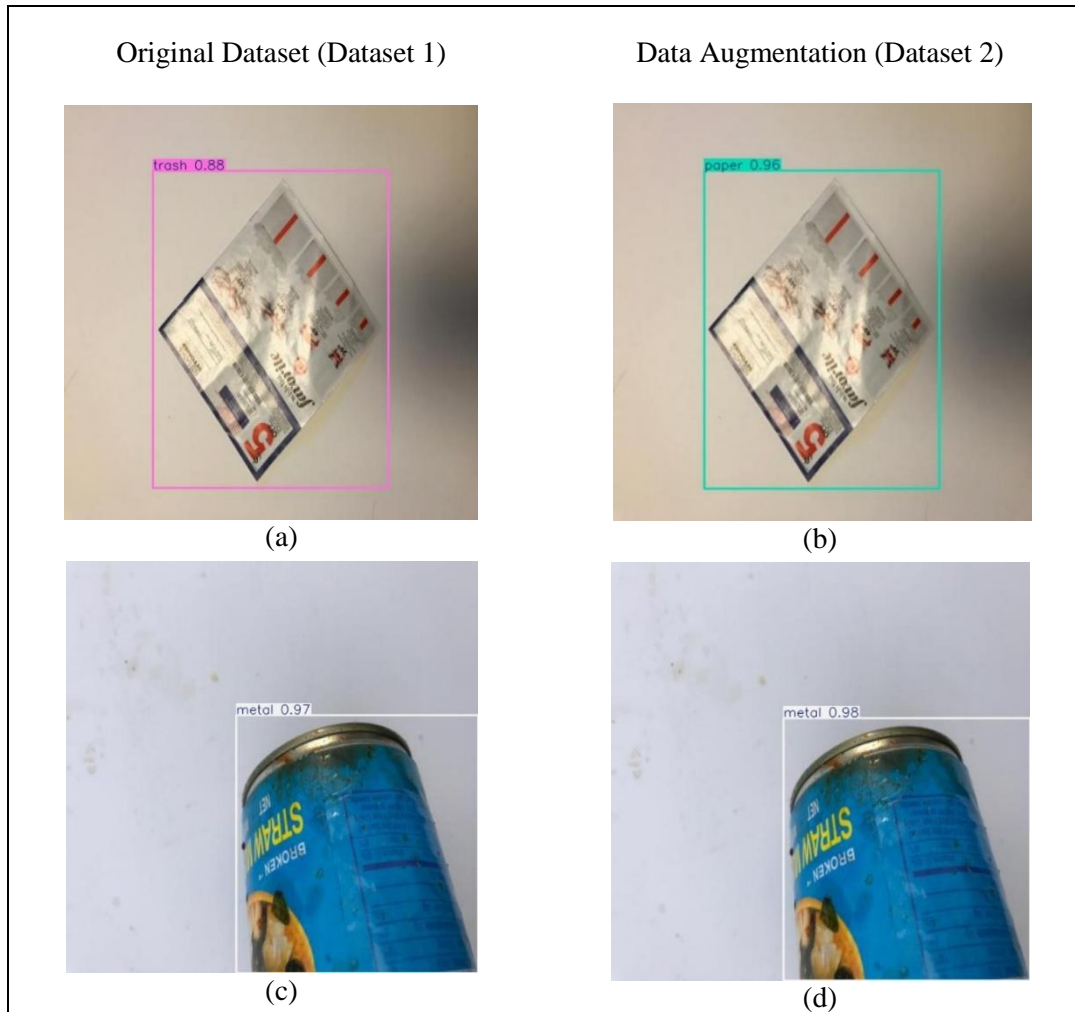


Table 3

Comparison of Deep Learning Methods from Previous Research in Object Detection

Author	Methodology	Accuracy
Awe et al. (2017)	Faster R-CNN	68.30%
Knowles et al. (2018)	OscarNet (CNN)	88.42%
Y. Liu et al. (2018)	YOLOv2	89.2%
Vo et al. (2019)	DNN-TC	94%
Ruiz et al. (2019)	Inception-ResNet	88.60%
Ye et al. (2020)	YOLO-VAE	69.70%
Chen et al. (2020)	YOLOv4	64%
Kumar et al. (2021)	YOLOv3	92.50%
Flores (2023)	YOLOv5	90.2%
Experiment	YOLOv8 using Dataset 2	96.5%
Proposed Method	YOLOv11 using Dataset 1	97.4%
Proposed Method	YOLOv11 using Dataset 2	97.8%

Table 3, proposed by Flores (2023), represents a comparison of several previous studies with this research. Some methods, such as Faster R-CNN, ResNet, and previous YOLO versions, will be compared with the performance of YOLOv11. The YOLOv11 model proposed in this research outperforms other approaches using the same dataset, as indicated by the accuracy comparison results of different approaches from earlier research. Research conducted by Awe et al. (2017) achieved a total of 68.30% accuracy using Faster R-CNN as the method. Research conducted by Knowles et al. (2018) using a CNN-based OscarNet as the method achieved a total of 88.42% accuracy. Meanwhile, the YOLOv2 method has been used by Y. Liu et al. (2018) and achieved a total of 89.2% accuracy.

Additionally, Vo et al. (2019) demonstrated performance improvement with a 94% accuracy rate using DNN-TC as the method, and Ruiz et al. (2019) achieved a total accuracy rate of 88.60% using Inception-ResNet as the method. However, other approaches, such as YOLO-VAE used by Ye et al. (2020) and YOLOv4 used by Chen et al. (2020), achieved only 68.70% and 64% accuracy, respectively. This indicates that not all of YOLO's developments perform consistently across all situations. Besides that, research conducted by Flores (2023) achieved a total of 90.2% accuracy using YOLOv5 as the method, while Kumar et al. (2021) achieved a total of 92.50% accuracy using YOLOv3 as the method, which was slightly better compared to YOLOv3 used by Flores (2023).

As a comparison, the YOLOv11 method used in this research achieved a total accuracy of 97.4% on the original dataset, which increased slightly to 97.8% accuracy due to the data augmentation technique. This result indicates that YOLOv11 is capable of highly precise object detection and classification, even surpassing all of the earlier methods shown in Table 3, including earlier versions of YOLO like YOLO-VAE, YOLOv2, YOLOv3, YOLOv4, and YOLOv5. The model's improved efficiency was mostly due to the combination of a more modern, effective architecture and enhanced data variety and quality through the data augmentation technique. As a result, YOLOv11 shows great potential for implementation in object detection tasks, particularly when it comes to garbage type classification, like in this research. It indicates that models were successfully trained and developed, resulting in an excellent performance.

The YOLOv11 models show potential in object detection tasks due to their speed and accuracy. However, there are still some difficulties when applied to garbage detection and classification. The difficulties are limitations in references and previous research of the YOLOv11 model, which was relatively new, limitations in computing resources and platforms usage while using Google Colab and Roboflow, and issues related to the quality and variety of the datasets that impact the training results of the models.

One of the first challenges is the lack of references and previous research on YOLOv11, a version of the recently released You Only Look Once (YOLO) model. It results in a lack of references or previous research on its architecture and performance, leading to a reference limitation. This is inversely correlated to YOLOv8, which has already been used in many research projects from various studies and fields. Researchers typically learn about the YOLOv11 model through tutorials and platforms like GitHub, gaining knowledge and insights to address the issue of limited references.

Limited computing resources and platform usage became issues that were challenging to resolve during the research process. Resources with usage restrictions for non-subscription users, such as Google Colab, can certainly disrupt the research process. Researchers often encounter situations during data training where the process times out due to memory allocation limits set by Google Colab or Roboflow. It needs to restart the process, particularly if the trained model has not been saved. The computational

requirements of larger YOLOv11 models present another limitation in this research. Larger variants typically perform better and have higher accuracy, but the model also demands a lot more computing resources. Due to these constraints, the smaller version of the YOLOv11 pretrained model was used for training and evaluation.

Even though this smaller model produces good results, it might not fully utilise the architecture's potential, especially in situations requiring more intricate object representations or deeper feature extraction. Similarly, Roboflow has a usage limit for non-subscription users, which restricts the quantity of images that can be uploaded and labelled. If the usage limit is exceeded, the augmentation process is also limited, which certainly impacts the range of datasets used for the training process. Certainly, the resource limitation significantly impacts the time efficiency of this research process, as researchers must wait for the process to begin after the usage limit is reset. The training process is usually done in the morning or afternoon to adjust the usage limit of Google Colab. Researchers must pay close attention during this process, as a prolonged period of inactivity from users may trigger a timeout and disconnect the runtime. To avoid problems that might disrupt the training's overall progress, the researcher has to stay fully present. Researchers gain valuable insights from this computing resource and platform usage limitation, including how to manage resources wisely and prevent unnecessary task utilisation.

The quality and variety of the dataset are crucial to this research. As a result, this problem has a significant effect on the overall model's detection performance. Low-resolution or ambiguous images might make extraction more difficult. Paper, glass, plastic, metal, cardboard, and trash are the six categories or classes of data used to support this research. Different camera angles, similar data textures, and shifting lighting can all result in inconsistent image quality. When the shadows appear on an image due to improper lighting, the labelling process becomes more cautious to avoid including the shadows in the bounding box. These factors certainly have an impact on the model's ability to detect and classify objects because different categories, like paper and cardboard or glass and plastic, which have similar characteristics, textures, and colours, are difficult to distinguish. It is also crucial to consider the variation that exists in the dataset, as it directly impacts the dataset's quality. The model can only identify conditions that match the available objects in the dataset if there is a lack of variation in the dataset. As a result, the model will struggle to identify and might even fail to identify objects in various states, such as bent cardboard, broken glass, folded plastic, and folded paper.

Additionally, there is one category or class in this dataset that is not on the same level of quantity as the other categories. Trash is the category which only contains about 100 photos. As a result, categories with fewer instances are predicted to be less accurate, and the model is biased towards categories with more instances. It not only affects the model's performance during the training process, but the quality and variety of these datasets also impact the model's ability to produce accurate predictions in real-world situations. The quality and variety of the datasets that will be used must be carefully considered. The developed YOLOv11 model will be able to perform precise and accurate detection in various situations and environmental conditions, provided the image quality is high and the dataset contains sufficient variations.

CONCLUSION

Based on the research that has been done, the YOLOv11 model used for object detection shows great potential when used to classify various categories of various types of garbage, including cardboard, glass, metal, paper, plastic, and trash. The model can attain an excellent level of accuracy, as evidenced

by its high success rate in making correct predictions. It indicates YOLOv11 can recognise and differentiate the key features of each category well. YOLOv11 is a groundbreaking innovation in real-time object detection, boasting a complex profile of significant advantages and fundamental disadvantages that warrant careful evaluation before application in practice or research. On the other hand, its improved training techniques and architectural improvements result in measurable improvements in accuracy and speed of inference compared to its previous versions, making it highly suitable for garbage sorting and recycling by allowing the development of advanced autonomous garbage management systems that are capable of accurately and instantly identifying a wide range of wasted materials. It will open up possibilities for automated and more effective robotic sorting that can work more accurately than the conventional manual techniques, ultimately improving resource recovery and decreasing landfill garbage.

This research also highlights some aspects that still need to be improved. Some misclassification occurs quite frequently. It indicates that there are similarities and characteristics between categories, such as cardboard and paper, or plastic and glass, which is one of the primary problems. For example, white cardboard is sometimes mistaken for paper, indicating that the model still struggles to distinguish between objects with similar shapes or colours. The model's performance in trash category classification is still not optimal due to the lack of images of trash. In addition, the quality and variety of objects must be improved. Models trained on datasets with limited variety often struggle because they cannot learn the broader and more diverse visual characteristics of the object category, resulting in decreased performance when encountering new objects. To address this issue, the researcher is increasing the diversity of the training data through the data augmentation process. The model's performance has significantly improved, as shown by the results, and the model can now identify objects more reliably and accurately. Even with a small amount of training data, the value of mAP50 from both experiments shows satisfactory performance. It supports the conclusion that the model's capacity to identify the characteristics of an object is significantly impacted by the quality and variety of the dataset, resulting from the data augmentation process.

However, when compared to other object detection models, the architecture of YOLOv11 showed higher accuracy and light weight. The shorter computation time makes the model more likely to be implemented on devices with limited resources. Strategies such as scheduling training time and optimising resource and platform usage have effectively addressed the limitations of limited GPU memory on platforms like Google Colab and Roboflow, thereby disrupting the training process and pre-processing data during the development phase. Despite facing various challenges, the YOLOv11 model proved to be effective in detecting and classifying garbage objects. This research also shows that YOLOv11 worked well in both datasets with satisfactory mAP50 values. It is believed that this research will contribute to developing an automated garbage management system in the future that is capable of identifying and classifying any kind of garbage accurately.

ACKNOWLEDGMENT

This research is supported by the Vice-Rector of Research, Innovation, and Entrepreneurship at Satya Wacana Christian University under Grant No. 122/SPK-PF/RIK/08/2025.

REFERENCES

- Adedeji, O., & Wang, Z. (2019). Intelligent waste classification system using deep learning convolutional neural network. *Procedia Manufacturing*, 35, 607–612. <https://doi.org/10.1016/j.promfg.2019.05.086>
- Ahmed, A., Farahani, S., Ahmed, S., Abbas, W., & Mehboob, S. (2024). VisionFusion (Vis-Fusion): Integrating CNNs and Vision Transformers for superior object detection. *European Journal of Theoretical and Applied Sciences*, 2(6), 851–860. [https://doi.org/10.59324/ejtas.2024.2\(6\).75](https://doi.org/10.59324/ejtas.2024.2(6).75)
- Ali, M. L., & Zhang, Z. (2024). The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection. *Computers*, 13(12), 336. <https://doi.org/10.20944/preprints202410.1785.v1>
- Alif, M. A. R. (2024). YOLOv11 for vehicle detection: Advancements, performance, and applications in intelligent transportation systems. <http://arxiv.org/abs/2410.22898>
- Aradea, Rianto, & Mubarak, H. (2024). Recognition of organic waste objects based on vision systems using attention convolutional neural networks models. *Scientific Journal of Informatics*, 11(3), 595–608. <https://doi.org/10.15294/sji.v11i3.6494>
- Aryantiningsih, D. S., Fitriani, I. M., Novita, R., Parlin, W., & Safitri, F. E. (2024). Pemberdayaan masyarakat mengenai pemilahan sampah dan pengolahan sampah organik menjadi pupuk kompos. *Jurnal Medika: Medika*, 3(2), 72-76.
- Arip, A. A. S., Sazali, N., Kadirgama, K., Jamaludin, A. S., Turan, F. M., & Ab. Razak, N. (2024). Object detection for safety attire using YOLO (You Only Look Once). *Journal of Advanced Research in Applied Mechanics*, 113(1), 37–51. <https://doi.org/10.37934/aram.113.1.3751>
- Arvio, Y., Kusuma, D. T., & Sangadji, I. B. (2024). Inorganic waste detection application using smart computing technology with YOLOv8 method. *Sinkron*, 8(4), 2389–2396. <https://doi.org/10.33395/sinkron.v8i4.14117>
- Awe, O., Mengistu, R., & Sreedhar, V. (2017). *Final report: Smart Trash Net: Waste localization and classification*. <https://cs229.stanford.edu/proj2017/final-reports/5226723.pdf>
- Bahagia, G. A., & Akbar, M. (2024). Klasifikasi sampah organik dan anorganik menggunakan metode convolutional neural network (CNN). *Jurnal Mahasiswa Teknik Informatika (JATI)*, 8(5), 10349-10355.
- Banduka, N., Tomić, K., Živadinović, J., & Mladineo, M. (2024). *Enhancing automated defect detection and classification using YOLOv11: Case study of finished leather industry*. <https://doi.org/10.20944/preprints202411.0239.v1>
- Cao, Y., & Yongke Sun. (2024). The research on the application of deep learning in education. *IETI Transactions on Data Analysis and Forecasting (ITDAF)*, 2(3), 4–11. <https://doi.org/10.3991/itdaf.v2i3.51413>
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). *End-to-end object detection with transformers*. <http://arxiv.org/abs/2005.12872>
- Chen, Q., & Xiong, Q. (2020). Garbage classification detection based on improved YOLOV4. *Journal of Computer and Communications*, 08(12), 285–294. <https://doi.org/10.4236/jcc.2020.812023>
- Emmanuel, N., Oberösterreich, F., & Lawal, M. M. (2024). Deep learning in computer vision: From object detection to automobile vehicles. *International Research Journal of Engineering and Technology (IRJET)*, 11(8), 301-312
- Fan, B., & Song, X. (2024). A review of deep learning based object detection algorithms. *Journal of Engineering Research and Reports*, 26(11), 88–99. <https://doi.org/10.9734/jerr/2024/v26i111316>

- Fitriana, F., Dalimunthe, N. P., Apriani, F., Priyansah, S., Arif, M., & Irwan, A. G. (2024). Penyuluhan pendidikan peduli lingkungan melalui pemilihan jenis sampah bagi siswa sekolah dasar. *Jurnal Masyarakat Mandiri (JMM)*, 8(1), 1277. <https://doi.org/10.31764/jmm.v8i1.20932>
- Flores, E. J. C. (2023). A YOLOv5s Model for classification of garbage. *Southeast Asian Journal of Science and Technology*, 8(1), 9-18.
- Ismail, G., Sidik, S., Adriana, A., & Nugraha, I. F. (2023). Identifikasi sumber, jenis dan kuantitas sampah domestik (studi kasus perusahaan jasa kecil). *Jurnal Syntax Admiration*, 4(10), 2092–2103. <https://doi.org/10.46799/jsa.v4i10.948>
- Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2024). Evaluating the evolution of YOLO (You Only Look Once) models: A comprehensive benchmark study of YOLO11 and its predecessors. <http://arxiv.org/abs/2411.00201>
- Kang, Z., Yang, J., Li, G., & Zhang, Z. (2020). An automatic garbage classification system based on deep learning. *IEEE Access*, 8, 140019–140029. <https://doi.org/10.1109/ACCESS.2020.3010496>
- Khanam, R., Asghar, T., & Hussain, M. (2025). Comparative performance evaluation of YOLOv5, YOLOv8, and YOLOv11 for solar panel defect detection. *Solar*, 5(1). <https://doi.org/10.3390/solar5010006>
- Khanam, R., & Hussain, M. (2024). YOLOv11: An overview of the key architectural enhancements. <http://arxiv.org/abs/2410.17725>
- Knowles, J., Kennedy, S., & Kennedy, Tom. (2018). OscarNet: Using transfer learning to classify disposable waste. *CS230 Report: Deep Learning. Stanford University, CA, Winter*.
- Kuang, E. Z., Bhandari, K. R., & Gao, J. (2024). Optimising waste management with advanced object detection for garbage classification. <http://arxiv.org/abs/2410.09975>
- Kumar, S., Yadav, D., Gupta, H., Verma, O. P., Ansari, I. A., & Ahn, C. W. (2021). A novel YOLOv3 algorithm-based deep learning approach for waste segregation: Towards smart waste management. *Electronics (Switzerland)*, 10(1), 1–20. <https://doi.org/10.3390/electronics10010014>
- Liu, H. (2024). The application of deep learning in the field of robotics. *Applied and Computational Engineering*, 93(1), 98–104. <https://doi.org/10.54254/2755-2721/93/20240928>
- Liu, Y., Ge, Z., Lv, G., & Wang, S. (2018). Research on automatic garbage detection system based on deep learning and narrowband Internet of Things. *Journal of Physics: Conference Series*, 1069(1). <https://doi.org/10.1088/1742-6596/1069/1/012032>
- Putri, S. S. I., Shindy, M., & Raden, A. Z. M. (2022). Sosialisasi klasifikasi sampah melalui mural. *Darma Cendekia*, 1(2), 114-123. <https://doi.org/10.60012/dc.v1i2.16>
- Ruiz, V., Sánchez, Á., Vélez, J. F., & Raducanu, B. (2019, May). Automatic image-based waste classification. In *International work-conference on the interplay between natural and artificial computation* (pp. 422-431). Cham: Springer International Publishing.
- Sapkota, R., Meng, Z., Churuvija, M., Du, X., Ma, Z., & Karkee, M. (2024). Comprehensive performance evaluation of YOLOv12, YOLOv11, YOLOv10, YOLOv9 and YOLOv8 on detecting and counting fruitlet in complex orchard environments. *arXiv preprint arXiv:2407.12040*. <https://doi.org/10.36227/techrxiv.172954111.18265256/v1>
- Sapkota et al. (2025). YOLO advances to its genesis: A decadal and comprehensive review of the You Only Look Once (YOLO) series. *Artificial Intelligence Review*, 58(9). <https://doi.org/10.1007/s10462-025-11253-3>

- Singh, R., Goyal, S., Agarwal, S., & Upadhyay, S. (2024). Evaluating the performance of ensembled YOLOv8 variants in smart parking applications for vehicle detection and license plate recognition under varying lighting conditions. <https://doi.org/10.20944/preprints202404.1216.v1>
- Vo, A. H., Hoang Son, L., Vo, M. T., & Le, T. (2019). A novel framework for trash classification using deep transfer learning. *IEEE Access*, 7, 178631–178639. <https://doi.org/10.1109/ACCESS.2019.2959033>
- Vrsalovic, I., Lerga, J., & Ivasic-Kos, M. (2025). A system for real-time detection of abandoned luggage. *Sensors*, 25(9). <https://doi.org/10.3390/s25092872>
- Yang, H., Chen, L., Chen, M., Ma, Z., Deng, F., Li, M., & Li, X. (2019). Tender tea shoots recognition and positioning for picking robot using improved YOLO-V3 model. *IEEE Access*, 7, 180998–181011. <https://doi.org/10.1109/ACCESS.2019.2958614>
- Yang, M., & Thung, G. (2016). Classification of trash for recyclability status. *CS229 Project Report*, 2016(1), 3.
- Ye, A., Pang, B., Jin, Y., & Cui, J. (2020). A YOLO-based neural network with VAE for intelligent garbage detection and classification. 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence, 1–7. <https://doi.org/10.1145/3446132.3446400>
- Zhang, D. (2024). Overview of object detection based on deep learning. *Applied and Computational Engineering*, 104(1), 40–46. <https://doi.org/10.54254/2755-2721/104/20241062>
- Zhang, M., Ye, S., Zhao, S., Wang, W., & Xie, C. (2025). Pear object detection in complex orchard environment based on improved YOLO11. *Symmetry*, 17(2), 255. <https://doi.org/10.3390/sym17020255>
- Zhou, Y., Lin, L., & Wang, T. (2024). Garbage classification detection system based on the YOLOv8 algorithm. *AIP Advances*, 14(12). <https://doi.org/10.1063/5.0244795>