Classification of Recyclable Waste Using Deep Learning: A Comparison of Yolo Models

Slamet Riyadi*, Annisa Divayu Andriyani, Ahmad Musthafa Masyhur

Information Technology Department, Universitas Muhammadiyah Yogyakarta, Yogyakarta 55183, Indonesia

Corresponding Author Email: riyadi@umy.ac.id

ABSTRACT

This study aimed to enhance the effectiveness of waste management systems to reduce harm to the environment and promote sustainability. The main objective of this study was to assess the capability of YOLOv5 and YOLOv8 to detect potential recyclable waste materials. YOLOv5 and YOLOv8 deep learning architectures, primarily focused on modern features, recognize and handle recyclable waste materials. The study application was based on the waste recycling plant dataset (WaRP), which was created to facilitate this platform's modification for garbage classification machine learning. YOLOv5 has been remarkably successful at waste detection because it uses a multi-scale detection system. The most memorable achievement of this study is the radical improvement in its detection accuracy and speed of performing the detection task seen in YOLOv8. YOLOv8 is a super-ordinate model that outperforms its predecessors' detection speed and resolution. The baud rate of the various object classes is unequal across object classes. Detection accuracy must be higher, especially for categories with below-average output performance. Precision and recall values for YOLOv5 and YOLOv8 are 0.478 and 0.569 and 0.442 and 0.513, respectively. Approximately 40.6% and 54.6% of the mAP50-IOU and 42.6% values outperform the range offered by YOLOv5. The improved performance of YOLOv8 compared to YOLOv5 demonstrates this platform's potential to enable a more accurate and timely waste management system while conserving the environment.

1. INTRODUCTION

The increase in awareness concerning the environment's preservation and proper use of resources has seen growing attention to sustainable approaches in terms of minimizing harm done to the environment. Developing effective waste management systems is significant since there is an increasing demand for sustainable approaches that safeguard the surroundings [1]. Applying some of these technologies would be especially effective when implemented using technologies such as YOLOv5 and YOLOv8, with a promise to monitor, recognize, and control recyclable garbage. Although previous research applied similar technology, significant gaps provided space for further study to increase the efficiency and effectiveness of these systems [2]. Issues with world garbage production and its negative environmental impact further underline the need for developing waste management systems. Specialists claimed that without appropriate care for it, the amount of trash produced worldwide would increase by up to 70% in 2025. This is thus a means to solve many environmental issues, such as pollution or the exhaustion of natural resources. This calls for innovative solutions with the latest techniques in garbage separation, recycling, and proper management to ensure a sustainable future [3].

Throughout history, technology has been crucial in advancing waste management procedures. Advanced technologies have progressively enhanced conventional approaches, such as sensors, robotics, and artificial intelligence (AI). Deep learning models such as YOLO (You Only Look Once) have become influential tools for tasks involving the detection and classification of objects in this situation. YOLO models are highly suitable for real-time applications because of their exceptional speed and precision, which makes them perfect for recognizing recyclable items in various waste streams [4].

This study analyzes the implementation of YOLOv5 and YOLOv8 technologies in recyclable trash management. Emphasis is placed on the identification and efficient management of waste, elaborating on potential responses that these technologies offer to resolve current problems related to waste management. YOLOv5 and YOLOv8 are two generations from the YOLO family with rare technical enhancements that make them more effective. For instance, YOLOv8 features more advanced network structures, improved methods for training, and more robust algorithms for detecting objects compared to YOLOv5. These enhancements improve accuracy, more complete detection, and improved performance. This is particularly evident when measured by the mean average precision (mAP) metric, a standard way to evaluate the effectiveness of garbage sorting systems. This research, however, aims to clarify the benefits of implementing YOLOv5 and YOLOv8 technologies for developing an improved waste management system in terms of speed, accuracy, and flexibility in sorting different types of garbage.

The research intends to improve the accuracy of separating

recyclable trash by developing models using YOLOv5 and YOLOv8. It also attempts to fine-tune the model and assess effectiveness through much experimentation. The model evaluation uses the WaRP dataset, a good repository of images showing different recyclable items. The wide range of variations and comprehensiveness of information within the dataset makes for an excellent benchmark to assess the model's potential. The study considers several scenarios for using the models in real waste management settings and, hence, yields practical insights into the potential benefits of these models [5].

As such, the study's outcomes will significantly improve waste management systems' efficiency and eco-sustainability. The study's results may serve as an affluent theoretical base for further development and sustainable ways of garbage management. This paper, therefore, is expected to stir a novel development on the topic and, at the same time, contribute to the efforts of global sustainability by reporting the performances of YOLOv5 and YOLOv8 in the detection and classification of recyclable materials [6].

2. LITERATURE REVIEW

2.1 Deep learning

Deep learning is a set of techniques that enables computer models to learn multi-level representations of data using multiple processing layers to learn features automatically. In essence, the models used for deep learning utilize artificial neural networks. This research has been done since the 1980s and was recently revived with the advent of faster computers [7]. Deep learning and its complementary concept is a new machine-learning field utilized increasingly with GPU acceleration development. Deep learning has been widely used in text mining, spam detection, video recommendation systems, image classification, and multimedia concept retrieval. However, deep learning has the following problems: there may be no sufficient train data, the volume of train data might be unbalanced, and there is no guarantee of scalability [8]. Further studies have utilized the application of deep learning in other fields, in particular, software engineering and mental disorders. A new systematic study was conducted to capture the most recent deep-learning methods in the prediction process [9].

2.2 Object detection

Literature studies on object detection have confirmed that deep learning is the most advanced method for object detection. An example is the implementation of deep learning by various algorithms to undertake multiple object detection, including the Scale Invariant Feature Transform, Convolutional Neural Networks, Region-Based Convolutional Neural Networks, Fast Region-Based Convolutional Networks, Faster Region-Based Convolutional Networks, You Only Look Once, and Single Shot Detector [10]. Another survey mentioned that many modern object detection algorithms are applied using deep learning, and the advantages and disadvantages of each object detector were reviewed carefully. Object detection also has wide applications and is critical to vision-based software systems [11].

The literature review also reveals that object detection refers to identifying instances of the classes specified, for example,

faces, cars, or trees, among others, in images or video. Unlike classification, object detection identifies multiple objects and locations within an image. Information usually returned from the object detector would be a list of detected objects and the object class information, probability score, and object coordinates. Object detection is also used in other areas, such as construction: it helps to enhance machine vision understanding of construction and safety activities in detecting objects [12].

2.3 YOLO

You Only Look Once (YOLO) is one of the deep learning models first developed in 2015 by Joseph Redmond. YOLO is one approach to object detection in real-time. The system processes detection with a custom classifier or localizer—an application of the model at various positions and scales on an image. YOLO has emerged as one of the most popular methodologies used in object detection. In addition, YOLO was trained to perform good and consistent person detections in frontal and asymmetric views [13].

YOLO is considered a real-time object detection approach. Detection has been achieved through the custom localizer or classifier. The model is applied at various positions and scales of the image. YOLO is one of the prevalent methodologies in object detection. Additionally, YOLO was trained to detect the position of persons either frontally or asymmetrically, which provides person detection results that are strong and reliable [14].

Besides, YOLO has also found applications such as license plate detection, pose detection, and text detection. Thus, YOLO has been one of the mainstream approaches in real-time object detection and video surveillance applications. The fact that YOLO is good at detecting objects tells a lot. Notably, this technology sprints, processing the images remarkably. It is indeed well-matched for applications with timely detection, such as in vehicles and surveillance. YOLO achieves high precision by detecting and classifying objects in the image, ensuring the fast processing of information. The system will be versatile and able to identify and classify objects suitable for different applications, including people and vehicles.

YOLO was designed with an end-to-end training approach that provides a simple training process with optimal performance in object detection. YOLO can estimate shifts at different scales, meaning it can still pinpoint objects of any size in an image. This continued active support from the community has made the YOLO one of the state-of-the-art, continuously improving with new architectural advancements, techniques in training, and deployment. This ongoing progress further helps strengthen its effectiveness in tasks related to object detection [15, 16].

3. METHODOLOGY

This paper implemented both YOLOv5 and YOLOv8 architectures in deep learning. The researcher applied deep learning to analyze photos from waste recycling plants. The objective of the study was to compare the performance of both architectures. The YOLOv5 architecture emphasizes feature extraction effectiveness and introduces a combination of Convolution, Batch Normalization, and SiLU activation modules into this process. On the other hand, YOLOv8 improved with features like the CSPDarknet53 backbone and

a PAN-FPN structure, enhancing feature integration and increasing detection accuracy.

3.1 Dataset

The WaRP dataset, standing for waste Recycling Plant Dataset, was a methodically created collection of annotated photos taken at an industrial waste sorting facility. It is a vital resource for machine learning and computer vision applications, specifically garbage sorting and recycling. It specialized in categorizing items into five main sections: Bottles, Carton, Detergent, Canisters, and Cans, out of which every section comprised 28 distinct variations. Apart from these are some data characteristics, such as various plastic bottles and cardboard types divided into paperboard and corrugated board. Furthermore, packages marked with the suffix "-full" indicate that their bottles contain air, unlike empty bottles.

WaRP was meticulously categorized into five primary sections: Bottles, Carton, Detergent, Canisters, and Cans, each encompassing 28 variations. There were 17 plastic bottles with the prefix "bottle" and three distinct categories of bottles with the prefix "glass." Cardboards were categorized into two main kinds, namely paperboards and corrugated boards. The collection consisted of products branded with the suffix "-full" to signify that these bottles contained air, differentiating them from empty bottles.

The WaRP dataset's distinguishing characteristic was its authenticity and depiction of demanding real-world situations. The images in this series faithfully depicted the challenges that emerged from the regular overlapping of objects, dramatic modifications, or unfavorable lighting circumstances. Incorporating this genuine component is crucial for the teaching and comprehensive assessment of machine learning models, particularly those intended for classifying waste under suboptimal conditions. Tags enclose the user's input. The primary element of Warp, WaRP-D, comprises a substantial quantity of photos utilized for training and validation. The dataset consists of 2452 training photographs that are crucial for the development of reliable waste sorting algorithms. Additionally, 522 validation images were used to evaluate the performance. The Warp photos provide a detailed depiction of rubbish found at recycling facilities, with a resolution of 1920x1080 pixels in high quality. The dataset's exceptional resolution renders it very appropriate for numerous computer vision and deep learning applications, particularly those that prioritize precise detection and efficient garbage sorting [17].

3.2 YOLOv5

YOLOv5 is an improved object detection model and has achieved state-of-the-art performance on many tasks, such as trash identification. Figure 1 shows the model architecture of YOLOv5. YOLOv5 offers several advantages that can be very useful for identifying trash recycling facilities. The core part of YOLOv5 corresponds to the feature extraction process and includes a focus module, CBS (Convolution, Batch Normalization, SiLU activation) modules, and C3_n (CBS, BottleneckCSP1, Concatenation) modules; "n" is the number of layers. Such architecture allows the model to effectively gather features at all scales and resolutions essential for precise object detection with their different sizes and forms [18].

In YOLOv5, the neck, typically a transition stage, connects the backbone and head. Its vital job is to enhance features

retrieved by the backbone further to improve the model in detecting objects of various sizes [19]. The YOLOv5 head performs predictions for object classes and regresses the object boundaries. YOLOv5 uses a decoupled head, which increases the performance of detection [20]. The reason behind making this decoupled head is that YOLOv5 is designed with a multi-scale detection methodology to detect objects of different scales. This function could be helpful for waste detection tasks since objects can vary both in size and shape [21].
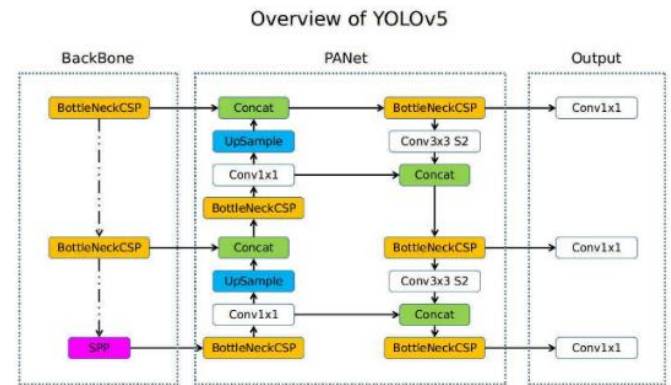

Figure 1. YOLOv5 architecture [22]

YOLOv5 has been used in waste detection tasks to detect construction waste and garbage functions. This model has been further enhanced using techniques like CBAM-CSPDarknet53, SimSPPF (Simplified SPPF), and ODConv, which has reflected the model YOLOv5-OCDS at the performance of garbage detection at its best rate [23].

3.3 YOLOv8

The YOLOv8, belonging to the YOLO series, outperformed the older versions, YOLOv5 and YOLOv7, in accuracy and speed of detection. As for the network architecture, the backbone of YOLOv8 mostly stayed the same as YOLOv5, only replacing C3 modules with CSP. The widely used SPPF module was installed after the backbone, ensuring high-accuracy measurements for an extensive range of magnitudes. It could fuse the PAN-FPN features effectively at different scale levels within the neck area. Besides, the Neck module consisted of many C2f modules, up-sampling layers, and the unique form of the head to further enhance the accuracy. The overall architecture of YOLOv8 is shown in Figure 2.
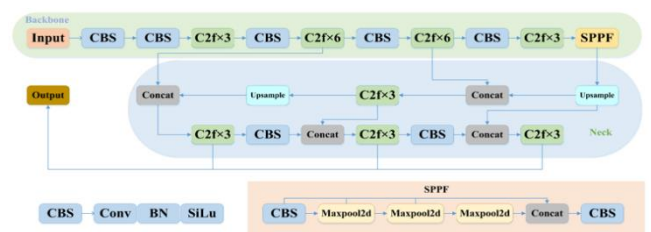

Figure 2. YOLOv8 architecture [20]

Modified CSPDarknet53 was the essential part of YOLOv8 and was named the backbone. The multi-scale approach was done by upscaling the input features into five independent scales represented as B1 to B5. In this respect, and towards enhancing content transmission while maintaining a lean architecture, the first CSP module was replaced with C2f,

made up of a chain of direct connections. CBS consisted of a convolution operation followed by batch normalization, further activation with SiLU, and, ultimately, the output. At the heart of the architecture, the SPPF module was employed to encourage the scaling of elastic production. It reduced the computation burden and latency by consecutively connecting three max-pooling layers [21].

For the YOLOv8 network, the PAN-FPN was used in the neck and based on the PANet layout. This architecture helps merge features and extract their locations, simplifying the PAN framework by eliminating the convolution operation after the sampling. The YOLOv8 model achieves optimum effectiveness and, at the same time, preserves its extraordinary levels of performance [24]. The PAN-FPN structure embeds the PAN and FPN frameworks' separate feature scales, for instance, P4-P5 and N4-N5. PAN-FPN was the feature fusion through top-down and bottom-up deep semantic and surface location data methods to increase feature variety and make it more comprehensive [25].

The head in YOLOv8's detection module—specifically called the detection part—chiefly came with a split-head structure that employed multiple branches for object classification and bounding box regression when making predictions. The definition uses three loss functions: BCE Loss for classification, DFL for bounding box regression, and CIOU for bounding box regression. The model implemented a Decoupling Design, improving detection precision while speeding up the model convergence. We further adopted a Task Specifier to achieve online dynamic sample selection [26].

4. RESULT AND DISCUSSION

In one elaborate experiment, the efficiency of YOLOv5 and YOLOv8 models was measured to determine the performance of the same datasets. Both models presented similar results during training, which is quite interesting to note. Moreover, it must be brought into focus that YOLOv8 showed better performance in each feature.

4.1 YOLOv5

The researcher conducted the training process using the YOLOv5 model over 100 epochs, as shown in Figure 3. The figure illustrates several metrics that track the model's performance during the training and validation phases. The train/box_loss, train/obj_loss, and train/cls_loss curves indicate a steady decrease in loss values, demonstrating the model's improved learning and convergence over time. The validation losses, as shown by val/box_loss, val/obj_loss, and val/cls_loss, similarly decrease, suggesting that the model generalizes well to unseen data.

The metrics/precision and metrics/recall plots indicate the model's ability to detect and classify objects accurately. Precision measures the proportion of correctly identified objects out of all objects identified, while recall measures the proportion of actual objects correctly identified by the model. Precision and recall values fluctuate in the early epochs, indicating variability in the model's learning process. However, as training progresses, these metrics stabilize, demonstrating a balance between the ability to find all relevant objects (recall) and to avoid false positives (precision). The metrics/mAP_0.5 and metrics/mAP_0.5:0.95 curves reflect the mean average

precision, a comprehensive measure of the model's performance. The steady rise in these values suggests an improvement in the model's detection accuracy across different intersection-over-union (IoU) thresholds.
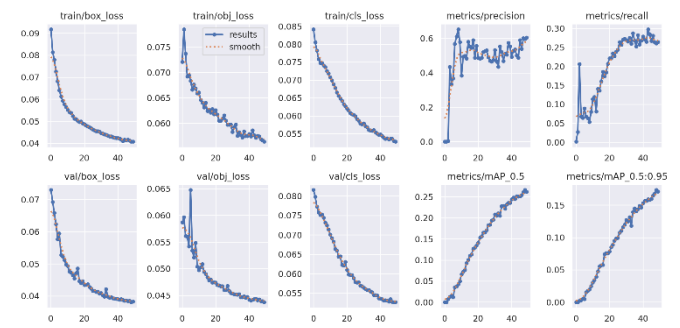

Figure 3. Training result of YOLOv5 model

Figure 4 displays the confusion matrix for the YOLOv5 model, which assesses the model's performance in classifying objects into various categories. The matrix highlights how well the model predicts each class and identifies areas of misclassification.

The overall metrics—precision: 0.478, recall: 0.442, mAP50: 0.406, and mAP50-95: 0.275—suggest that the model performs reasonably well, but noticeable variation across different classes exists. Some classes, such as 'bottle-green-full,' demonstrate high recall (0.912) and decent precision (0.493), indicating that the model effectively identifies these objects with relatively few errors.

In contrast, the model struggles with classes like 'detergent-box' and 'detergent-transparent,' where precision and recall are near zero. This indicates significant challenges in correctly identifying these items, possibly due to class imbalance, similar visual features, or inadequate training data for these categories.

The confusion matrix reveals the need for further refinement of the YOLOv5 model, especially for classes with low detection accuracy. Potential improvements could include rebalancing the dataset, fine-tuning the model's parameters, or employing more sophisticated data augmentation techniques to enhance overall classification accuracy.
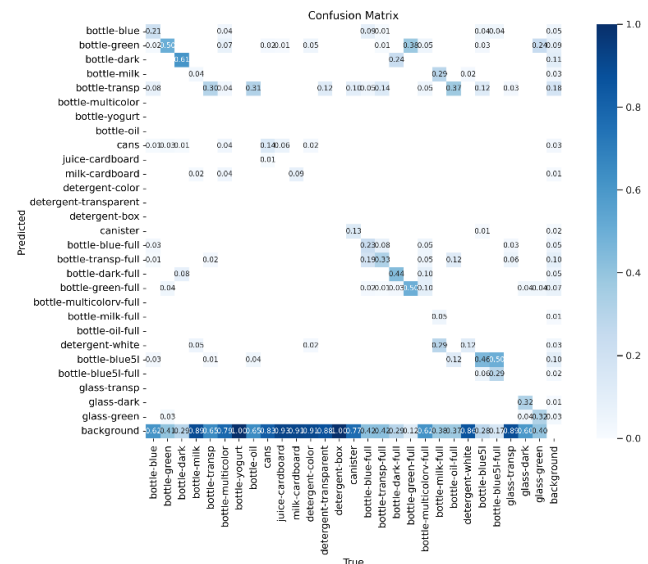

Figure 4. Confusion matrix of YOLOv5 model

Table 1 presents the specific metrics of precision, recall, mAP50, and Map50-95 values, explicitly focusing on the 28 feature classes emphasized throughout the training phases.

These data offer a comprehensive assessment of the YOLOv5 model's performance in these specific classes.
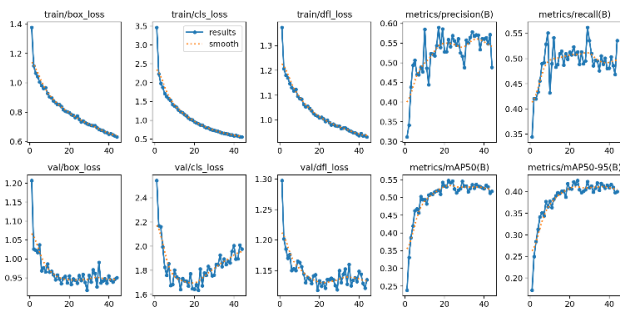
Table 1. Model summary training phase on YOLOv5 model

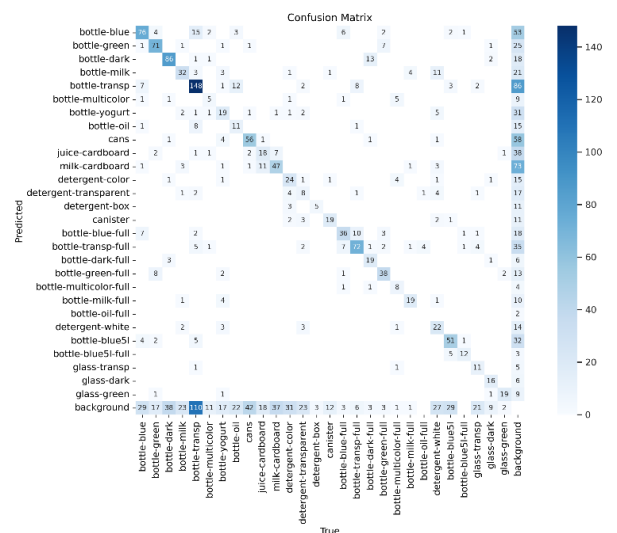| Class	Precision	Recall	mAP50	mAP50-95
all	0.478	0.442	0.406	0.275
bottle-blue	0.382	0.577	0.586	0.32
bottle-green	0.422	0.703	0.57	0.387
bottle-dark	0.556	0.768	0.766	0.522
bottle-milk	0.253	0.333	0.205	0.139
bottle-transp	0.413	0.466	0.38	0.247
bottle-multicolor	1	0	0.0638	0.397
bottle-yogurt	0.228	0.19	0.127	0.0821
bottle-oil	0.266	0.146	0.14	0.108
cans	0.341	0.296	0.29	0.183
juice-cardboard	0.351	0.221	0.21	0.123
milk-cardboard	0.333	0.489	0.376	0.24
detergent-color	0.35	0.233	0.231	0.174
detergent-transparent	0.317	0.0488	0.0744	0.0513
detergent-box	1	0	0.0481	0.0325
canister	0.47	0.333	0.256	0.181
bottle-blue-full	0.407	0.698	0.599	0.401
bottle-transp-full	0.515	0.62	0.576	0.409
bottle-dark-full	0.555	0.794	0.739	0.532
bottle-green-full	0.493	0.912	0.851	0.64
bottle-ulticolorv-full	0.555	0.333	0.471	0.342
bottle-milk-full	0.475	0.857	0.679	0.534
bottle-oil-full	1	0	0.0394	0.0278
detergent-white	0.142	0.215	0.124	0.0859
bottle-blue51	0.479	0.714	0.598	0.394
bottle-blue51-full	0.477	0.792	0.739	0.613
glass-transp	0.428	0.306	0.295	0.144
glass-dark	0.519	0.6	0.667	0.301
glass-green	0.65	0.744	0.755	0.437

4.2 YOLOv8

The researcher starts the training process using the YOLOv8 model in 100 epochs, as shown in Figure 5. The confusion matrix of the YOLOv8 model can be seen on Figure 6.

Figure 5. Training result of YOLOv8 model

The given data shows the precision, recall, mAP50, and mAP50-95 values for different object detection classes. As per 'all' categories, the overall performance is precision, 0.569; recall, 0.513; mAP50, 0.547; mAP50-95, 0.426. Some engaging individual class performances were 'bottle-dark-full,' which has a high precision of 0.853, and 'bottle-oil-full,' with a high recall of 0.846. The mAP50-95 metrics balance thresholds of confidence variation reflect how smoothly the model performs. However, for some classes like 'bottle-multicolor' and 'detergent-transparent,' precision and recall

values are on the lower side. This is further analyzed with the changes to be made correctly to get better precision and recall of robust object detection across different categories.

Figure 6. Confusion matrix of YOLOv8 model

Table 2 presents comprehensive data on precision, recall, mAP50, and Map50-95 metrics. This data is specifically related to the 28 feature classes that were the primary focus during the training phases. This data provides a thorough evaluation of the performance of the YOLOv8 model in these categories.

Table 2. Model summary training phase on YOLOv8 model

Class	Precision	Recall	mAP50	mAP50-95
all	0.569	0.513	0.547	0.426
bottle-blue	0.541	0.584	0.591	0.448
bottle-green	0.708	0.667	0.738	0.571
bottle-dark	0.754	0.638	0.77	0.591
bottle-milk	0.519	0.508	0.51	0.432
bottle-transp	0.622	0.453	0.553	0.413
bottle-multicolor	0.269	0.217	0.226	0.185
bottle-yogurt	0.353	0.316	0.319	0.263
bottle-oil	0.391	0.271	0.3	0.236
cans	0.504	0.515	0.507	0.36
juice-cardboard	0.304	0.396	0.308	0.237
milk-cardboard	0.342	0.44	0.369	0.29
detergent-color	0.553	0.343	0.408	0.312
detergent-transparent	0.298	0.205	0.17	0.142
detergent-box	0.327	0.625	0.638	0.497
canister	0.535	0.576	0.553	0.5
bottle-blue-full	0.589	0.782	0.694	0.547
bottle-transp-full	0.633	0.776	0.766	0.637
bottle-dark-full	0.853	0.612	0.785	0.647
bottle-green-full	0.726	0.764	0.798	0.637
bottle-ulticolorv-full	0.714	0.499	0.557	0.466
bottle-milk-full	0.695	0.846	0.791	0.672
bottle-oil-full	0.768	0.2	0.336	0.236
detergent-white	0.576	0.299	0.417	0.344
bottle-blue51	0.596	0.536	0.619	0.495
bottle-blue51-full	0.656	0.715	0.724	0.575
glass-transp	0.72	0.275	0.45	0.34
glass-dark	0.712	0.478	0.596	0.365
glass-green	0.665	0.833	0.812	0.493

4.3 Result comparison

Performance comparison between the object detections of YOLOv5 and YOLOv8 models are recorded in Table 3. The model summary for the training phases of YOLOv5 and YOLOv8 on different classes of object detections is provided in the data.

Table 3. Object detection performance comparison between YOLOv5 and YOLOv8 models

Model	Precision	Recall	mAP50	mAP50-95
YOLOv5	0.478	0.442	0.406	0.275
YOLOv8	0.569	0.513	0.547	0.426

In the YOLOv5 training, the model demonstrates an overall precision of 0.478, recall of 0.442, mAP50 of 0.406, and mAP50-95 of 0.275. The performances are significantly noteworthy for classes like 'bottle-dark-full' and 'bottle-green-full,' which have high precision, recall, and mAP scores, indicating effective detection of objects. However, the difficulties can be noticed in certain classes, for example, 'bottle-multicolor' and 'detergent-box,' whose values of both precision and recall are hampered, hence suggesting difficulty in the identification of those objects with accuracy.

In contrast, YOLOv8 performs much better in total precision, at 0.569; total recall, at 0.513; mAP50, at 0.547; and mAP50-95, at 0.426. Marked improvements in other classes are shown under precision, recall, and mAP values for 'bottle-dark-full' and 'bottle-green-full.' Comparative results indicate the progress from YOLOv5 to YOLOv8, which guarantees higher ability and performance, especially in object detection. More detailed analysis and attention to specific class-level improvements will direct focused fine-tuning for optimized

precision and recall of the object detection model.

It is now essential to understand that deep learning is hounded by problems of imbalanced data and poor scalability, which gives us the ground to understand its limitations and makes some recent advances and suggestions significant in this respect—advances like data augmentation, transfer learning, and the use of synthetic datasets. Further, the addition of distributed computing and better algorithm efficiency can deal with many more scaling problems so that models can process far larger datasets and do much more complicated tasks.

5. CONCLUSIONS

This study investigates the application of YOLOv5 and YOLOv8 technologies in managing recyclable garbage. The YOLOv5 model exhibited satisfactory performance in detecting objects but encountered challenges in identifying other categories. The YOLOv8 model showed substantial enhancements in precision, recall, mAP50, and mAP50-95 compared to YOLOv5, namely in its ability to detect 'bottle-dark-full' and 'bottle-green-full' objects. The study emphasizes the capacity of sophisticated artificial intelligence technologies to enhance waste management systems and increase environmental sustainability. The YOLOv5 model attained a precision of 0.478, a recall of 0.442, and a mean average precision at 50% intersection over union (mAP50) of 0.406 and 0.275, correspondingly. The YOLOv8 model exhibited exceptional precision, recall, mAP50, and mAP50-95, affirming its outstanding performance in trash detection. The study offers practical insights into the potential utilization of YOLOv5 and YOLOv8 in real-world trash management settings. It also acts as a source of inspiration for future

research in sustainable waste management techniques.

REFERENCES

[1] Abdu, H., Noor, M.H.M. (2022). A survey on waste detection and classification using deep learning. IEEE Access, 10: 128151-128165. https://doi.org/10.1109/ACCESS.2022.3226682

[2] Sary, I.P., Andromeda, S., Armin, E.U. (2023). Performance comparison of YOLOv5 and YOLOv8 architectures in human detection using aerial images. Ultima Computing: Jurnal Sistem Komputer, 15(1): 8-13. https://doi.org/10.31937/sk.v15i1.3204

[3] Kaza, S., Yao, L., Bhada-Tata, P., Van Woerden, F. (2018). What a waste 2.0: A global snapshot of solid waste management to 2050. World Bank Publications. https://doi.org/10.1596/978-1-4648-1329-0

[4] Hameed, S., Rahman, F.U., Amir, M., Babrawala, Z.U.A., Hassan, M.U. (2021). IoT based waste management system. Journal of Robotics and Mechanical Engineering, 1(1). https://doi.org/10.53996/2770-4122.jrme.1000105

[5] Choi, J., Lim, B., Yoo, Y. (2023). Advancing plastic waste classification and recycling efficiency: Integrating image sensors and deep learning algorithms. Applied Sciences, 13(18): 10224. https://doi.org/10.3390/app131810224

[6] Chu, Y., Huang, C., Xie, X., Tan, B., Kamal, S., Xiong, X. (2018). Multilayer hybrid deep-learning method for waste classification and recycling. Computational Intelligence and Neuroscience, 2018(1): 5060857. https://doi.org/10.1155/2018/5060857

[7] Boukerche, A., Hou, Z. (2021). Object detection using deep learning methods in traffic scenarios. ACM Computing Surveys (CSUR), 54(2): 1-35. https://doi.org/10.1145/3434398

[8] Rasidi, A.I., Pasaribu, Y.A.H., Ziqri, A., Adhinata, F.D. (2022). Klasifikasi sampah organik dan non-organik menggunakan convolutional neural network. Jurnal Teknik Informatika dan Sistem Informasi, 8(1): 142-149. https://doi.org/10.28932/jutisi.v8i1.4314

[9] Neu, D.A., Lahann, J., Fettke, P. (2022). A systematic literature review on state-of-the-art deep learning methods for process prediction. Artificial Intelligence Review, 55(2): 801-827. https://doi.org/10.1007/s10462-021-09960-8

[10] Pathak, A.R., Pandey, M., Rautaray, S. (2018). Application of deep learning for object detection. Procedia Computer Science, 132: 1706-1717. https://doi.org/10.1016/j.procs.2018.05.144

[11] Schwalm, D. (2023). Detecting objects on IP camera video with Tensorflow and OpenCV. Analytics Vidhya. https://medium.com/analytics-vidhya/detecting-objects-on-ip-camera-video-with-tensorflow-and-opencv-e2c25297a75a, accessed on Nov. 6, 2023.

[12] Shrigandhi, M.N., Gengaje, S.R. (2022). Systematic literature review on object detection methods at construction sites. In International Conference on Expert Clouds and Applications, pp. 709-724. https://doi.org/10.1007/978-981-99-1745-7_52

[13] Muhlashin, M.N.I., Stefanie, A. (2023). Klasifikasi Penyakit Mata Berdasarkan Citra Fundus Menggunakan YOLO V8. JATI (Jurnal Mahasiswa Teknik Informatika), 7(2): 1363-1368. https://doi.org/10.36040/jati.v7i2.6927

[14] Afonso, M.H., Teixeira, E.H., Cruz, M.R., Aquino, G.P., Boas, E.C.V. (2023). Vehicle and plate detection for intelligent transport systems: Performance evaluation of models YOLOv5 and YOLOv8. In 2023 IEEE International Conference on Computing (ICOCO), pp. 328-333. https://doi.org/10.13140/RG.2.2.11022.95042

[15] Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., Shen, H., Ren, J., Han, S., Ding, E., Wen, S. (2020). PP-YOLO: An effective and efficient implementation of object detector. arXiv preprint arXiv:2007.12099. https://doi.org/10.48550/arXiv.2007.12099

[16] Quach, L.D., Quoc, K.N., Quynh, A.N., Ngoc, H.T. (2023). Evaluating the effectiveness of YOLO models in different sized object detection and feature-based classification of small objects. Journal of Advances in Information Technology, 14(5): 907-917. https://doi.org/10.12720/jait.14.5.907-917

[17] "WaRP - Waste Recycling Plant Dataset." https://www.kaggle.com/datasets/parohod/warp-waste-recycling-plant-dataset, accessed on Dec. 25, 2023.

[18] Ghatkamble, R., Parameshachari, B.D., Pareek, P.K. (2022). YOLO network based intelligent municipal waste management in internet of things. In 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), pp. 1-10. https://doi.org/10.1109/ICERECT56837.2022.10060062

[19] Sun, Q., Zhang, X., Li, Y., Wang, J. (2023). YOLOv5-OCDS: An improved garbage detection model based on YOLOv5. Electronics, 12(16): 3403. https://doi.org/10.3390/electronics12163403

[20] Tamin, O., Moung, E.G., Dargham, J.A., Yahya, F., Farzamnia, A., Sia, F., Naim, N.F.M., Angeline, L. (2023). On-shore plastic waste detection with YOLOv5 and RGB-near-infrared fusion: A state-of-the-art solution for accurate and efficient environmental monitoring. Big Data and Cognitive Computing, 7(2): 103. https://doi.org/10.3390/bdcc7020103

[21] Ma, M., Tam, V.W., Le, K.N., Li, W. (2020). Challenges in current construction and demolition waste recycling: A China study. Waste Management, 118: 610-625. https://doi.org/10.1016/j.wasman.2020.09.030

[22] YOLOv5 model architecture overview. (2024). https://github.com/ultralytics/yolov5/issues/4518

[23] Wu, Z., Zhang, D., Shao, Y., Zhang, X., Zhang, X., Feng, Y., Cui, P. (2021). Using YOLOv5 for garbage classification. In 2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), pp. 35-38. https://doi.org/10.1109/PRAI53619.2021.9550790

[24] Karwa, R.R., Gupta, S.R. (2022). Automated hybrid deep neural network model for fake news identification and classification in social networks. Journal of Integrated Science and Technology, 10(2): 110-119.

[25] Wang, Y., Zhang, X. (2018). Autonomous garbage detection for intelligent urban management. In MATEC

Web of Conferences, 232: 01056. https://doi.org/10.1051/matecconf/201823201056

[26] Lou, H., Duan, X., Guo, J., Liu, H., Gu, J., Bi, L., Chen, H. (2023). DC-YOLOv8: Small-size object detection algorithm based on camera sensor. Electronics, 12(10): 2323. https://doi.org/10.3390/electronics12102323