Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 24 July 2026
Revised 20 Aug 2026
Accepted 05 Sep 2026

Natural Resources for Human Health 2026, Vol. 6 (22s): 156–168

KEYWORDS:
YOLOv26, hard-negative mining, higher resolution input, intelligent waste-segregation systems

Towards Robust Intelligent Waste Segregation: A Diagnostic and Mitigation YOLOv26 Framework Using Hard-Negative Mining and Higher Input Resolution

Supriya[1], Ashutosh Shukla[2], Vineet Kumar[1], Shishant[1], Sonit Raj[1], Shivam Kumar Prajapati[1], Anmol Gupta[1]

[1]Department of Artificial Intelligence and Machine Learning, College of Smart Computing, COER University, Roorkee, Uttarakhand, India
[2]Department of Electrical Engineering, Haridwar University, Roorkee, Uttrakhand,India

ABSTRACT: Effective waste segregation is essential for improving recycling efficiency and reducing the environmental impact of improper waste disposal. Manual waste sorting is time consuming and inefficient, thus, the need for automated vision-based waste detectors is needed. The aim of this study is to explore the use of YOLOv26 for automated detection and classification of waste objects with a custom-made dataset. The mentioned dataset consists of 1,979 annotated images, out of which 1,707 are training images, 192 are validation images and 80 are testing real world capture images that belong to 9 waste classes, which includes categories of hazardous, organic and inorganic waste. The analysis of the confusion-matrix for the classes showed that the main issue was not inter-class confusion but that significant confusion occurred between the waste objects and the background. This failure mode was tackled by using hard-negative mining techniques to bias the model against false-positive background patterns, and Higher Input Resolution that raises the number of input pixels from 640 × 640 to 960 × 960 to increase the ability to detect small objects and visually similar patterns. The confusion-matrix analysis showed that background confusion remained the main residual error after mitigation, although the mitigation improved precision and mAP. The final mAP@0.5 improved from 0.7140 to 0.7259, while mAP@0.5:0.95 increased from 0.5805 to 0.6048. Precision also improved from 0.7590 to 0.7796, indicating a reduction in false-positive detections. The results show that the proposed technique of diagnostic error analysis with targeted hard-negative mining and higher resolution input can improve the robustness of YOLOv26 for autonomous detection of waste and serve as a foundation for further development of intelligent waste-segregation systems.

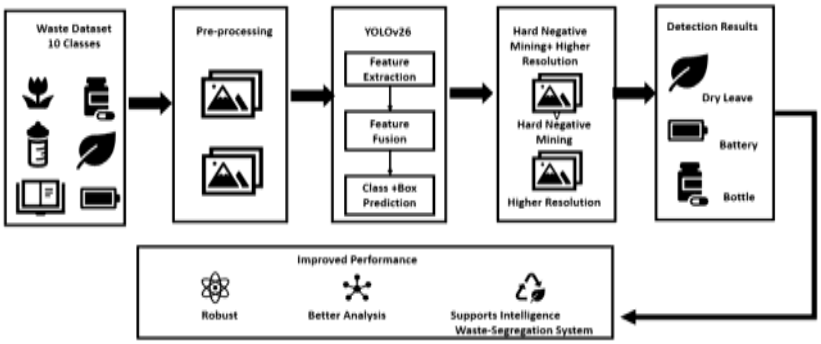


FIGURE 1. Graphical Abstract for the Proposed Model

Natural Resources for Human Health
eISSN: 2583-1194

1. INTRODUCTION

Waste management is a major problem in developing nations. Waste categorization is an essential and significant step. The practice of classifying or separating comparable types of garbage is known as waste classification. Fifty to seventy-five percent of the garbage is household waste [1]. Since organic garbage makes up the majority of home waste, efficient waste segregation reduces the amount of waste that ends up in landfills. Degradable garbage is different from non-biodegradable waste. Paper, textile, and agricultural waste make up the majority of organic or biodegradable garbage, whereas plastics, metals, rubber, glass, and other materials make up non-biodegradable waste. Electrical, hazardous, and medical waste were among the numerous forms of garbage that were not treated. On the other hand, trash output has recently grown significantly [2].

Inadequate waste management could have a detrimental impact on the environment. Sorting the garbage will therefore maximize optimization. In order to reduce the possibility of environmental harm, the quantity of recyclable materials should be recovered at the preliminary stage. This will enhance the general health of the local people in addition to being good for the environment. Thus, trash sorting ought to be carried out at the initial stage of waste management, which seeks to recycle as many items as possible and send less to landfills, potential for cross-contamination with other items. Unprofessional personnel isolate waste, which is inefficient, takes longer, and is less productive because a lot of time is lost. Once more, garbage has enormous economic worth once it is separated. The garbage is valuable if it is separated and used again

Currently, the waste collection and sorting process is still done manually by human at waste area, which takes a long time to detect organic and inorganic waste. The following stages will not be able to process metal, glass and toxic waste since it is not possible to transform these into electrical energy. However, waste having a high organic content is recommended for immediate biological waste treatment for producing methane gas. Plastic/Styrofoam waste needs to be pre-processed by the process of Drying, Shredding and Moulding into waste pellets.

The most recent technical advancements make their utilization a beneficial thing. Thus, the execution of Machine learning and artificial intelligence can produce a good solution to this concerning issue and to keep our surroundings a healthy place for everyone to live. [3]. In this research as shown in figure 1, the object detection system uses the YOLOv26 method of the Python Ultralytics library, which is the YOLO26, the latest Ultralytics release, real-time computer vision model, offering several improvements over the previous version. This study makes several contributions:

- This research explores evaluation of the performance of YOLOv26 for the accurate and efficient identification of different types of inorganic waste.
- A new dataset is created and this dataset is employed to train, validate, and test the proposed waste detection model based on YOLOv26.
- To uncover specific failure modes of the baseline model, a detailed class-wise confusion matrix analysis is performed, showing systematic confusion between the object classes and the background class; targeted mitigation strategies, such as hard-negative mining and raising the resolution of the input images, are implemented, with a quantitative improvement of the baseline model demonstrated through a before and after comparison.
- The developed dataset and experimental results, such as diagnostic error analysis and comparative evaluation, can be used as a reference in the future for studies related with deep learning for waste detection, classification and automatic waste segregation.

This study has four sections: **Introduction:** Where a general background on the relevance of segregation of waste and the potential application of computer vision to improve it. It also includes related work that is related to research topic and its suggested solution since it provides procedural overview of study which was necessary**. Materials and Methods**: Technical Approach - a methodology whereby a project is divided into small, sequential parts**. Results and Discussion**: All the experimentation done will be critically analysed in this section. **Conclusion**: Conclusion based on experimental analysis have been carried out this research section. It also includes recommendations that can be used to further enhance this research will be placed in this section.

There are a number of recent studies dedicated to the development of intelligent systems for accurate categorization and separation of waste in organic and inorganic categories. For achieving this goal, researchers have resorted to real-time image collection or to publicly available images and image repositories. These images are then pre-processed and classified using different machine learning and deep learning algorithms such as Artificial Neural Network (ANN), Convolutional Neural

Network (CNN) and You Only Look Once (YOLO) object detection models. These approaches are aimed at enhancing the accuracy and efficiency of waste detection and classification in various conditions of the environment. The main findings and approaches presented in these studies are summarized in Table 1

Table 1. Comparative analysis of existing waste-management approaches

Ref.	Year	Dataset	Classes	Model	Performance & Research Gap
[4]	2024	2,000 images; hybrid dataset	Plastic bottles, Plastic, Glass, Cans, Styrofoam	YOLOv8	Average accuracy: 98%. **Gap:** Further validation on diverse real-world waste conditions.
[5]	2025	Primary data	Organic, Inorganic	ResNet50	Accuracy: 97.5%; >99% with augmentation. **Gap:** Dataset expansion and evaluation with other architectures required.
[6]	2025	1,298 images; Kaggle & ImageNet	Metal, Organic, Paper/Cardboard, Plastic, Glass	YOLOv8	Accuracy: 88.05%. **Gap:** Dataset diversity and real-world generalization remain challenging.
[7]	2025	Public datasets; 1,500 images	Organic, Non-organic, Recyclable	YOLOv11, CNN, SSD, Faster R-CNN	Accuracy: 85.02%; F1: 85%. **Gap:** Limited real-world testing and dataset diversity.
[8]	2025	1,000 self-collected images	Organic, Inorganic, B3	YOLOv11 + IoT	mAP@50: 97.0%; mAP@50–95: 87.5%. **Gap:** Background interference and limited dataset diversity.
[9]	2026	2,326 images; TACO + RecyBat24 + Food Waste	Organic, Inorganic, B3, Residue	YOLOv11, YOLO26	YOLO26 mAP@50: 0.506; mAP@50–95: 0.393. **Gap:** Residue detection and real sorting-line validation remain challenging.
[10]	2024	TrashNet; 2,520 images	Cardboard, Glass, Metal, Paper, Plastic, Trash	YOLOv8n/s/m	Accuracy: 88.86–91.25%. **Gap:** Limited real-world diversity and generalization.
[11]	2023	2,442 images	7 waste categories	MobileYOLO	mAP = 90.9%; 80.5 FPS. **Gap:** Limited open-source waste datasets and large labelled-data requirements; semi-supervised learning is suggested.
[12]	2024	2,350 original → 5,872 augmented images	Biodegradable, Plastic, Paper, Metal	YOLOv8m	Precision = 0.94188, Recall = 0.96622, mAP@50 = 0.98252, mAP@50–95 = 0.92723. **Gap:** Computational limitations and lower performance at stricter IoU ranges.

[13]	2025	Custom image dataset	Biodegradable, Non-biodegradable	YOLOv8 + Arduino-based separator	Precision = 97.0%, Recall = 96.1%, F1 = 96.55%, mAP@0.5 = 98.25%. **Gap:** Limited lighting, occlusion and background diversity; binary classification and visual ambiguity remain challenges.

To summarize, current deep-learning models have shown good performance in waste classification and detection. But the diversity of data sets, background interference, changes in the environment, and the generalization of real-world restrictions remain. Accurate, efficient and powerful waste detection systems are still needed. The YOLO models have potential to enhance the automated waste segregation and recycling processes. Thus, future studies should address the improvement of the detection accuracy, model robustness, and applicability of models to practical systems in waste management.

2. MATERIALS AND METHODS

The methodology of this study consists of five stages, as depict in figure 2: preparing a dataset, training a baseline model, conducting a diagnostic error analysis, implementing targeted mitigation, and comparative evaluation. This staged design enables a systematic identification of the failure modes of the model and a quantification of the impact of the proposed mitigation strategies can be done as a direct before and after comparison.

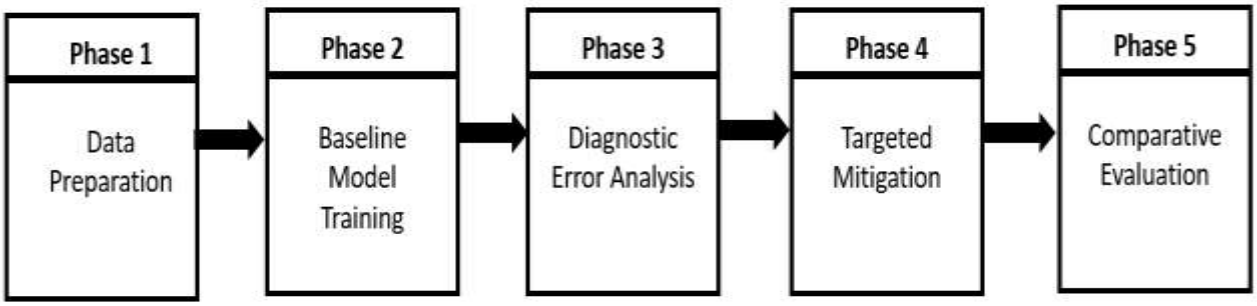


Figure 2. Phases for the proposed methodology

2.1 Dataset Preparation

We used a custom dataset including public repository dataset of 9 waste category and capture 80 real world images for test dataset class labels with images of each waste, each annotated with a label from the following regulatory classes: hazardous, i.e. Battery, Bugs Spray; organic, i.e. Dry Leave, Fruit Waste; and inorganic, i.e. Glass, Paper, Plastic Bag, Plastic Bottle, Tin. Bounding boxes were annotated in YOLO format and class labels were given for each object instance for each image. To train the images split into a training set, a validation set, and testing set. There are 1707 images in training set, 192 images in validation set, and real time 80 images in test set.

Table 2. Few prominent existing waste-management dataset

Ref.	Dataset	Size	No. Classes	Image Acquisition	Annotation	Public / Private
[14]	WADE	2,167	10	Consumer smartphones	Bounding box + category label	To be released publicly upon acceptance
[15]	GlobalWaste Data (GWD) archive	89,807	14	Mixed across source datasets: manual field photography, web scraping.	Directory-based	Public
[16]	TrashNet	2,527	6	Publicly available	Single image-level class label	Public — Kaggle

[17]	Orgalid Waste	5,600	4	Publicly available	Single image-level class label	Public dataset
[18]	Garbage Dataset (GD)	12,259	10	Multi-source: mobile app captures, web scraping	Manual annotation	Public —Kaggle
[19]	TriCascade WasteImage	35,264	2 &9	Combined from four existing public datasets	Images labelled	Public

This table 2 above shows the details of five waste-image data sets used for waste classification and detection, including the number of images per set, the number of waste classes in each set. The majority of datasets are publicly accessible, and most use image level class labels, with WADE providing bounding-box object annotations. While these datasets can be used to train deep learning models, differences in acquisition settings, class diversity and annotation style can impact actual performance. It is therefore critical to test models on real images taken in real contexts in a variety of light conditions for assessing their robustness and generalisability.

The training and validation images (1,707 and 192 images, nine waste categories) were taken from the public Roboflow Universe dataset Organic & Inorganic Detection (version 3; created 25 July 2024; https://universe.roboflow.com/organic-and-inorganic-waste-detection/organic-inorganic-detection), released under the CC BY 4.0 license and exported in YOLO format. The annotations in the form of bounding boxes are included with the data. We cleaned up the data by applying our own train/validation split and included background images that were not labelled with any of the classes in the dataset for hard-negative mining (Section 2.4). We have used the 80 real-world test images, they were not used for training or validation. [20]

2.2 Baseline Model Training

We trained our baseline model as an object detector using YOLOv26 and using normal augmentations and an input resolution of 640 for 250 epochs. The optimizer used was AdamW, with a learning rate of lr 0.000769 , best weights were selected by the Ultralytics fitness criterion (weighted combination of mAP@0.5 and mAP@0.5:0.95) For training, the batch size was 16 and 250 epochs, and the input size of the baseline and proposed model was 640 and 960, respectively. Optimizer and learning rate were automatically chosen by Ultralytics framework (AdamW, 0.000769, 0.9, 0.0005, respectively). The weights of the best model was chosen using Ultralytics fitness measure on a validation set. Experiments were run with Ultralytics 8.4.192 and PyTorch 2.11.0. [21]

The overall performance of an object detector can be expressed as:

$$Iimg \longrightarrow Fx(I_img) \longrightarrow \{Bi,Pi,Ci\} \quad (1)$$

Where ,

Iimg is the input image,

Fx represents the deep neural network, and

Pi the output consists of bounding boxes Bi classes and

Ci confidence scores

For every object detected, the detector estimates a bounding box of the detected object in the form of a rectangle. There are different ways to represent the bounding box[22,23]:

$$Bi = (x,y,wid,hgt) \quad (2)$$

Where,

x and y are the centre coordinates of the bounding box ,

wid and hgt, the width and height of the box.

Intersection over Union (IoU) is used to quantify the overlap between a predicted bounding box and its ground-truth bounding box.

It is defined as:

$$IoU=(Area(Bi\cap Bigt))/(Area(Bi\cup Bigt)) \quad (3)$$

Where, Bi is the predicted bounding box and

Bigt is the ground-truth bounding box.

The value of IoU is in the range of 0 to 1:

IoU = 0: no overlap, IoU of 0.9 or more: Nearly complete overlap,

IoU = 1: identical boxes

The IoU metric is essential to the evaluation of object detection, as it has to take into account both if the correct class is being detected and if the detected region is being placed in the right location. The concept of the confidence is based on the probability of the object's presence and the quality of the predicted localization [24-27].

The simplified representation is:

$$Ci = Pi\ (Object) \times IoU \quad (4)$$

where ,

Pi (Object) is the predicted probability of the presence of an object.

2.3 Error Diagnosis by Confusion Matrix Analysis

Using Ultralytics' in-built validation pipeline, a class-wise confusion matrix was built with an IoU threshold of 0.5 and a confidence threshold of 0.25 to get a better understanding of the failure modes of the baseline model beyond just detection metrics. Confusion matrix (raw count and row normalized) was created and from these the precision, recall, F1-score and AP@0.5 were calculated per class. This analysis showed very few confusions between different object classes, but a high and bidirectional confusion between all object classes and the background class. The background regions were frequently detected as waste objects (false positives), and some objects were missed (false negatives). The Battery class had the lowest object-class recall.

2.4 Targeted Mitigation

Two targeted interventions were used, based on the failure pattern found during previous phase:

- **Hard-Negative Mining**: After training the baseline model, it was run on the training images, and images on which it produced confident false-positive detections (predicted boxes matching no ground-truth object) were selected as hard negatives. These images were added to the training set as background images with empty label files, so that the model learns to suppress these patterns. The mined negatives came from the training data only; no validation, test or real-world images were used. The model was then retrained with the same hyperparameters, except for the input size (960 instead of 640).

- **Increased Input Resolution:** To make it easier to see small objects, the input resolution was raised from 640×640 to 960×960 pixels, and this increased the detectability of the Battery class, which had the highest rate of background confusion among the object classes for the baseline model.

2.5 Comparative Evaluation

The performance of the retrained model was then tested with the same held-out test set of the baseline model, but with the same confusion-matrix construction procedure described in Phase 3. This allowed for a "before and after" comparison of this direct effect.

Natural Resources for Human Health
eISSN: 2583-1194

3. RESULTS AND DISCUSSION

In this research will evaluate the performance of YOLOv26 in identifying and detecting organic and inorganic waste and assist in accelerating the waste sorting process. To train the images split into a training set, a validation set, and testing set. There are 1707 images in training set, 192 images in validation set, and 80 images in test set. Individual images are labelled manually with the appropriate waste category, allowing them to be used for systematic assessment of the proposed models.

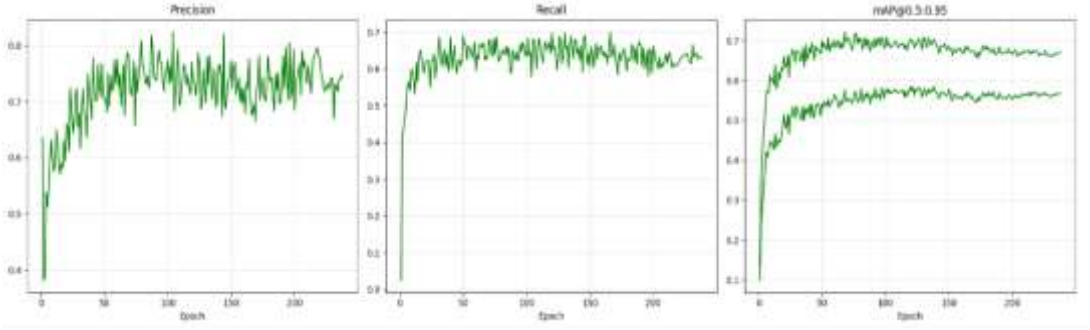
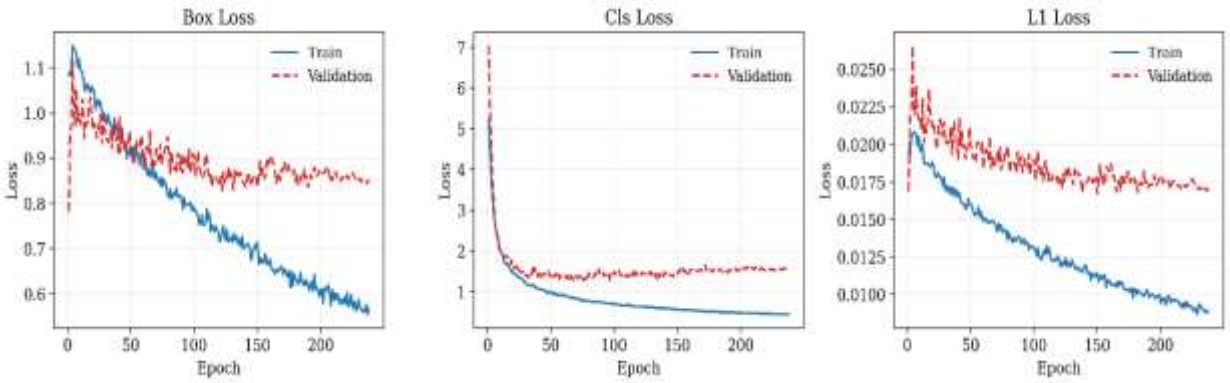


Figure 3. Evaluation Metric of the proposed model

Graphs shows in figure 3, a Class loss graph with a steady loss in the training class, and a lower loss at first but a higher loss in the validation class: with a small train–validation gap. This has indicated good learning during training with slight stability after about 50 epochs. The precision is fast increasing in the first epochs and it becomes stable around 0.70–0.78, which corresponds to good object classification. The recall increases rapidly to about 0.60-0.68 and is fairly constant, indicating that most relevant objects are being detected well. The mAP@0.5 is close to 0.70, which means that the performance of the detection is also high at IoU 0.5. As the model is tested for higher IoU (mAP@0.5:0.95), its performance stabilizes around 0.55-0.58. The overall metrics show a gradual decrease in their value after about 50-100 epochs, indicating that the model has stabilized and further training may lead to only marginal gains.

As depict in the figure 4 below, the training losses (Box, Classification, and L1) steadily decrease, indicating effective model learning. Validation losses decrease initially and then stabilize at a higher level, suggesting good convergence with some degree of overfitting after the early epochs. Both the training set and the validation set show Box Loss values decreasing over time, which is a sign of the model's improved ability to locate the bounding-boxes. The Box Loss value for the training set decreases steadily with time while the Box Loss value for the validation set decreases initially before levelling off around 0.85–0.90. The loss value of the class goes down significantly in the early epochs and slowly tends to converge to ~0.4 for train loss while the validation loss value stabilizes at 1.3-1.6. During training, L1 Loss always decreases and it's around 0.009 for training and 0.017 for validation. There is a slight difference between training and validation loss suggesting some degree of overfitting, especially after about 50-100 epochs. Overall, the training losses are decreasing, indicating effective learning, and the validation losses are relatively stable, suggesting that the model is not significantly improving through later epochs.



Natural Resources for Human Health
eISSN: 2583-1194

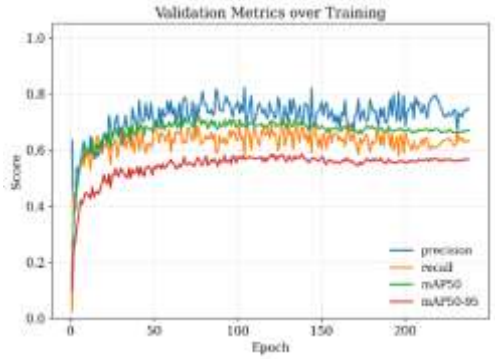


Figure 4. Losses of the proposed model

The confusion matrix shown in figure 5, depicts that the model is able to correctly classify most of the waste categories, with the values on the diagonal indicating high accuracy. Plastic Bottle has the most correct predictions (78), followed by Tin (35) and Dry Leave (30). The normalized matrix demonstrates good classification results, with normalized values of >0.60 for several classes. There are some mis-classifications between waste categories which are similar in appearance such as organic waste and inorganic waste, To mitigate this pattern classification failure, two enhancements are done. Hard-negative mining, which retrained the model on background-only images by identifying and re-annotating false positives as "no object," and increasing input resolution from 640×640 to 960×960 pixels to enhance the detectability of small objects, particularly the Battery class, which faced significant background confusion in the baseline model. As depict in figure 7 , the results prove that the model can make a good classification of the main waste classes with a comparatively low number of classification mistakes.

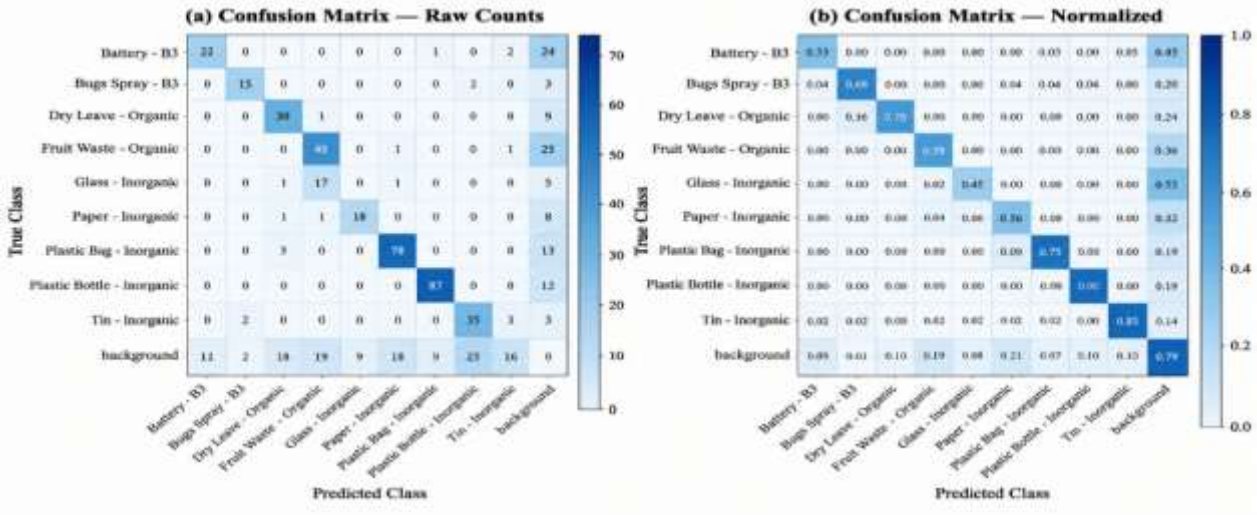


Figure 5. Confusion Matrix of YOLOv26

It is shown in the figure 6 that the confidence scores and bounding boxes are used to indicate the detection of objects for waste classification. It can classify various types of waste including plastic bag, coconut/palm waste, glass bottle and aluminium can waste. The model is able to accurately localize each waste object and predict its class with high confidence. The object detection model properly detects four types of waste or discarded things, which are paint box, flower garland, plastic bag, and coconut peel. The detected objects are depicted in green bounding boxes and the confidence scores range from 0.92 to 0.98 and depict the prediction confidence of the model. The model is able to accurately localize multiple paint boxes and individual waste items in various backgrounds. This indicates the possibilities of automated waste identification and segregation using object detection technique with YOLO.

Natural Resources for Human Health
eISSN: 2583-1194



Figure 6. confidence scores and bounding boxes of the Real world Images

3.4 Discussion

Confusion matrices shown in figure 7, are used to compare baseline model with the model used after Hard-Negative Mining for waste classification. The classes improved after hard-negative mining are Battery (0.42-0.53); Dry Leave (0.65-0.76); Tin (0.76-0.79) Reduced classification performance for some classes, such as Fruit Waste and Glass, however, highlights an opportunity for further improvement .As a whole, hard-negative mining enables the recognition of several hard classes, but it is difficult to get the background class right.

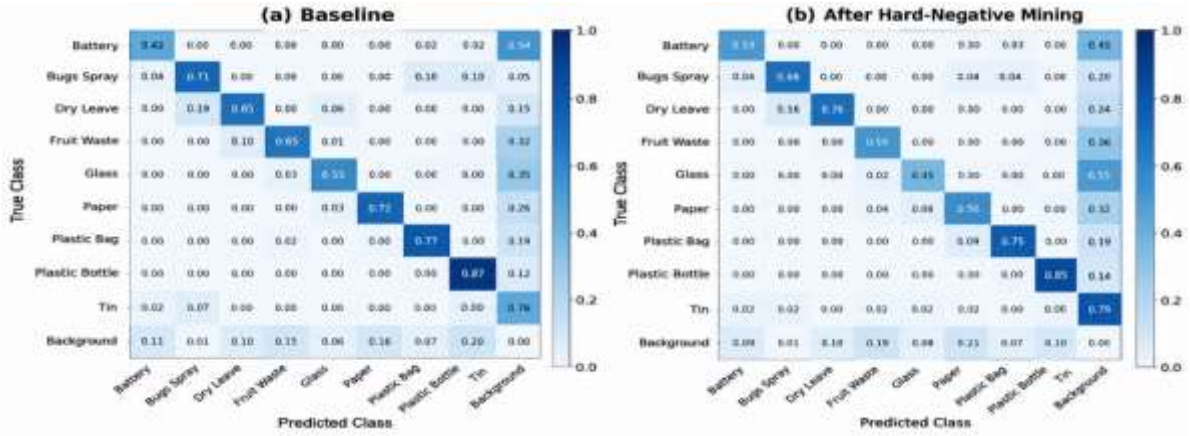


Figure 7. Confusion Matrix of YOLOv26 Baseline (a) and After HNM (b)

The experimental results demonstrate in table 4 , depicts the effectiveness of combining hard-negative mining with an increased input resolution from 640 × 640 to 960 × 960. The final mAP@0.5 improved from 0.7140 to 0.7259, while mAP@0.5:0.95 increased from 0.5805 to 0.6048. Precision also improved from 0.7590 to 0.7796, indicating a reduction in false-positive

detections. Although the final recall showed a slight variation (0.6440 vs. 0.6298), the improved precision, higher mAP values, indicate more effective object localization and classification.

Table 4. Comparative study of performance of proposed Baseline model and YOLOv26+ HNM+ HIR

Metric	YOLOv26+ HNM+ HIR	YOLOv26	Difference	Observation
Epochs	250	250	—	Same training duration
Precision	0.7796	0.7590	+0.0207	Higher precision in Proposed model
Recall	0.6298	0.6440	-0.0142	Slightly lower recall (precision–recall trade-off)
mAP@0.5	0.7259	0.7140	+0.0119	Better overall detection performance model
mAP@0.5:0.95	0.6048	0.5805	+0.0244	Better localization across IoU thresholds
Train Box Loss	0.6406	0.8042	−0.1636	Better bounding-box learning
Train Classification Loss	0.4141	0.6332	−0.2192	Better class discrimination
Validation Box Loss	0.7750	0.8777	−0.1027	Better localization on validation data
Validation Classification Loss	1.2766	1.3198	−0.0432	Slightly better validation classification
Best Precision	0.8136	0.7846	+0.0290	Higher peak precision
Best Recall	0.7239	0.7136	+0.0103	Higher peak recall in Proposed model
Best mAP@0.5	0.7406	0.7250	+0.0155	Higher peak detection performance
Best mAP@0.5:0.95	0.6088	0.5829	+0.0259	Better localization accuracy
Training Time	5746 s	7101 s	−1355 s	Proposed trained substantially faster

Table 5. Computational cost at 640 and 960 input sizes (Tesla T4, batch size 1)

Input size	Params	GFLOPs	Inference (ms)	Total (ms/img)	FPS	Peak GPU memory
640×640	2.51 M	5.91	9.48	12.54	79.8	77 MB
960×960	2.51 M	13.64	9.77	11.90	84.1	108 MB

Table 5 , The table above shows that the number of parameters remains the same (2.51 M), but the computational cost increases from 5.91 GFLOPs to 13.64 GFLOPs (≈2.3×) and peak GPU memory from 77 MB to 108 MB while measuring with a dummy 960×960 input, 200 runs after 20 warm-up runs.. The model inference time at 640 and 960 respectively was 9.48 and 9.77 millisecond per image (batch size 1) on Tesla T4 (GPU). This means that the model is able to infer at about 80–84 FPS, including pre- and post-processing. The training cost is incurred only when training with hard-negative mining. The proposed configuration is then valid for real-time application on this hardware, albeit with more FLOPs and memory might be significant on embedded hardware.

In figure 8, demonstrate the results and showcase that hard-negative mining combined with higher-resolution input improves the detector's ability to distinguish waste objects from challenging background regions, thereby enhancing the overall reliability of the waste detection system. The comparative detection performance of both models, baseline and enhanced detector with hard-negative mining and higher-resolution input, is shown. The results show that hard-negative mining is effective in improving the ability to learn to classify waste objects from similar background regions, since hard-negative samples are

emphasized during training. Moreover, the higher input resolution of 960 × 960 allows the detector to obtain more detailed spatial characteristics, textures, edges, and boundaries of the objects to be detected, mainly for small, overlapping, or partially hidden waste pieces. By combining these techniques, the rate of false detections can be lowered, and detection of the exact location of an object in complex scenarios can be enhanced, including cluttered surfaces and backgrounds that vary in texture and lighting. While several classes show improvement in the confusion matrix, the results are not consistent across all classes, suggesting that there is still room for improvement in optimizing the performance of each class. In total, these improvements increase the detector's capacity to detect waste under difficult conditions and help to create a more robust and efficient intelligent waste detection and segregation system.



Figure 8. comparative detection performance of both models, baseline and enhanced detector with hard-negative mining and higher-resolution input

4. CONCLUSION

This paper presented a fine-grained error analysis of a YOLOv26 detector trained on nine waste categories spanning organic, inorganic, and hazardous materials. The model showed strong discrimination between object classes, with minimal inter-class confusion. However, a systematic failure mode was identified: confusion between object classes and the background category, resulting in many background regions being detected as waste objects. This was attributed to small object size, textural similarity with background clutter, and insufficient background training samples.

To address this, hard-negative mining and an increased input resolution (640 to 960) were applied. After mitigation, precision and mAP improved, but confusion with the background class remained, indicating that further optimization is required. In the proposed YOLOv26 + HNM + HIR framework, the detection performance of the model is improved slightly, with a precision of 0.7796, mAP@0.5 of 0.7259, and mAP@0.5:0.95 of 0.6048. The model gains in performance of detection accuracy, localization of the bounding boxes and class discrimination compared to the baseline, which is backed by lower training and validation losses. Higher resolution input and hard-negative mining enable detection and discrimination of difficult waste items from complex background areas, thereby enhancing the accuracy and automation of waste detection and separation. However, there are still issues requiring further optimization and evaluation in the real world of background misclassification. These results confirm that class-wise diagnostic analysis, combined with targeted mitigation, is essential for building reliable automated waste-sorting systems.

Limitations. The reported results were obtained with a test set that is small in size (80 images), implying feasibility but not guaranteed performance in a variety of waste-management scenarios. In this study, only a single custom dataset was used and hard-negative mining and high input resolution were both used together. Hard-negative mining and high input resolution were not evaluated independently. Conclusions on deployment readiness should not be drawn because of the need for validation on larger multi-site datasets with different cameras, lighting conditions and waste composition and testing under live sorting-line conditions.

Future work will explore systematic confidence/IoU threshold tuning, expansion of background training data under varied conditions, evaluation on alternative detector architectures (e.g., Faster R-CNN), and deployment of the improved model in a real-time embedded waste-sorting prototype.

ACKNOWLEDGEMENTS

The authors acknowledge the technical assistance provided by the Department of AI & ML, COER University.

AUTHOR CONTRIBUTIONS

Supriya & Ashutosh Shukla: Conceptualization, Investigation.

Shishant & Sonit Raj: Methodology, Laboratory Analysis.

Shivam Kumar Prajapati & Anmol Gupta: Data Curation, Formal Analysis.

Vineet Kumar: Supervision, Writing—Review & Editing.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ETHICS STATEMENT

Ethical approval was not required for this study.

DATA AVAILABILITY STATEMENT

Data supporting the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] Gunawan, S. N., & Sudibyo, U. (2026). Development of an Intelligent Waste Identification System Based on the YOLOv11 Algorithm. Journal of Applied Informatics and Computing, 10(3), 2128-2138.

[2] Feriyanto, F., & Widiyanto, E. P. (2026). Implementation of Organic and Inorganic Waste Detection Modeling in School Waste Bins Using Yolov11. Jurnal Ilmiah Multidisiplin Indonesia (JIM-ID), 5(02), 279-288.

[3] Ling, L., Chen, Y., & Li, Z. (2026). WCD-YOLO: A waste classification detection model. Journal of Environmental Management, 398, 128601.

[4] Wadhwanwala, A. A. (2025). Heterogenous Waste Detection using YOLO (Doctoral dissertation, Dublin, National College of Ireland).

[5] Qinantha, I. K. M. C., Indrawan, I. G. A., Putra, I. P. S. U., Aristamy, I. G. A. A. M., & Willdahlia, A. G. (2025). Classification of organic and inorganic waste using ResNet50. Indonesian Journal of Data and Science, 6(2), 232-240.

[6] Arvio, Y., Kusuma, D. T., & Sangadji, I. B. (2024). Inorganic waste detection application using smart computing technology withYOLOv8 method. Sinkron: jurnal dan penelitian teknik informatika, 8(4), 2389-2396.

[7] Castañeda Meza, A. J., Lopez Haro, N. S. A., & Ornella Flores-Castañeda, R. (2025). Artificial Intelligence Based System for Sorting and Detection of Organic and Inorganic Waste. International Journal of Advanced Computer Science & Applications, 16(5), 87.

[8] Darisang, L. A., Dwisetio, S. K., Edbert, I. S., & Aulia, A. (2025). Real-time waste detection using YOLO, SSD, and Faster R-CNN integrated with CNN-based classification. In Proc. Int. Conf. on Research and Innovations in Information and Engineering Technology (RITECH) (pp. 76-83).

[9] Permana, M. F. J., Gani, J. C. R. B., Fauzan, N. A., & Adiwilaga, A. (2025). YOLOv11 Model as a Smart Solution for Waste Identification and Classification in Automated Waste Management System. Jurnal Ilmu Komputer dan Informasi, 18(2), 207-216.
[10] Vieri, D., Susanto, R., Purwanto, E. S., & Ario, M. K. (2024). Enhancing waste classification with YOLOv8 models for efficient and accurate sorting. Procedia Computer Science, 245, 889-895.
[11] Suman, M., & Arulanantham, G. (2023). Efficient Differentiation of Biodegradable and Non-Biodegradable Municipal Waste Using a Novel MobileYOLO Algorithm. Traitement du Signal, 40(5), 1833...
[12] Shrestha, S., Shrestha, S., Paudel, P., Gurung, S., Gairee, S., & Adhikari, S. (2024). Deep learning for waste management: Leveraging yolo for accurate waste classification. In 15th IOE Graduate Conference.
[13] Kumar, R. S. (2025). An automated waste separator for biodegradable and non-biodegradable materials using YOLOv8 and Arduino.
[14] Ijaz, M., Khan, S. U. R., Rehman, A. U., Asif, T., Vollmer, S., Dengel, A., & Asim, M. N. (2026). GlobalWasteData: A Large-Scale, Integrated Dataset for Robust Waste Classification and Environmental Monitoring. arXiv preprint arXiv:2602.07463.
[15] Shuvo, M. A., Farabi, A., Minhaz, M. A. A., Hasan, M., Khandaker, I., Shanto, I. K., & Kabir, M. N. (2026). WADE: A Reasoning-Annotated Benchmark for Multi-Instance Floating-Waste Grounding with Compact Vision-Language Models. arXiv preprint arXiv:2608.22950.
[16] Maleknya, K., Alizadehsani, R., Pedrammehr, S., & Shirini, K. (2026). A solid waste classification using reinforcement learning boosted by data augmentation and hyperparameter optimization. Scientific Reports.
[17] Hossen, M. M., Majid, M. E., Kashem, S. B. A., Khandakar, A., Nashbat, M., Ashraf, A., ... & Chowdhury, M. E. (2024). A reliable and robust deep learning model for effective recyclable waste classification. IEEe Access, 12, 13809-13821.
[18] Kunwar, S. (2026). The garbage dataset (GD): A multi-class image benchmark for automated waste segregation. arXiv preprint arXiv:2602.10500.
[19] Nahiduzzaman, M., Ahamed, M. F., Naznine, M., Karim, M. J., Kibria, H. B., Ayari, M. A., ... & Haider, J. (2025). An automated waste classification system using deep learning techniques: Toward efficient waste recycling and environmental sustainability. Knowledge-Based Systems, 310, 113028.
[20] Organic and Inorganic Waste Detection. (2024). Organic & Inorganic Detection (Version 3) [Data set]. Roboflow Universe. https://universe.roboflow.com/organic-and-inorganic-waste-detection/organic-inorganic-detection.
[21] Midigudla, R. S., Dichpally, T., Vallabhaneni, U., Wutla, Y., Sundaram, D. M., & Jayachandran, S. (2025). A comparative analysis of deep learning models for waste segregation: YOLOv8, EfficientDet, and Detectron 2. Multimedia Tools and Applications, 84(29), 35941-35964.
[22] Ramadhani, L. K., & Widyaningrum, B. N. (2026). Benchmarking YOLO26 Against YOLOv11 for Minority Class Waste Detection on an Augmented TACO Dataset. Journal of Applied Informatics and Computing, 10(4), 4100-4108.
[23] Kumar, M. G. V., Kumar, M., Rao, K. N., Rao, P. S., Tirumala, A., & Patnala, E. (2024, August). Advanced YOLO-based trash classification and recycling assistant for enhanced waste management and sustainability. In 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI) (pp. 1238-1246). IEEE.
[24] Schmidt-Mengin, M., Soulier, T., Hamzaoui, M., Yazdan-Panah, A., Bodini, B., Ayache, N., ... & Colliot, O. (2022). Online hard example mining vs. fixed oversampling strategy for segmentation of new multiple sclerosis lesions from longitudinal FLAIR MRI. Frontiers in Neuroscience, 16, 1004050.
[25] Manjunath, P. S. (2026). Agentic AI-Based Pest Monitoring Framework for Sustainable Agricultural Productivity and Resource Optimization. Natural Resources for Human Health, 6(16s), 740-750.
[26] Roopashree, S. (2026). A Unified Low-Latency YOLOv5–YOLOv8 Framework for Real-Time Iris–Pupil Detection in Biometric and Clinical Systems. Natural Resources for Human Health (NRFHH), 6(9s), 1430-1449.
[27] Wang, Z., Jiang, Q., Zhao, S., Feng, W., & Lin, W. (2023). Deep blind image quality assessment powered by online hard example mining. IEEE Transactions on Multimedia, 25, 4774-4784.