SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

# AI-Based Garbage Segmentation from Aerial Drone Images Using YOLOV8

Sugali Viswanatha Naik[1], Aravind Kumar Madam[2]

[1]Ph.D. Research Scholar, Department of ECE, B.E.S.T Innovation University, Andhra Pradesh, India
[2]Professor & Principal, Department of Electronics and Communication Engineering, West Godavari Institute of Science and Engineering, Prakashraopalem, Nallacherla Mandal, East Godavari District, Andhra Pradesh, India

**ABSTRACT**
The YOLOv8n-seg model has been used for AI garbage segmentation from drone images. The focus is on urban waste monitoring. The model achieved strong results for detection and segment performance with a box precision of 0.821, a recall of 0.656 and a mask mAP50 of 0.581 on the valid set. This metric reveals the capability of segmenting and localizing the trash correctly in the complex visual environment. The confusion matrix showed that the model performed very well in the classification of the background. However, there was one case where the background was incorrectly categorized as "garbage. This means that we need to further fine-tune the model so that both these classes are well-separated. Qualitative assessments demonstrated that in different aerial images, the garbage and other solid wastes were able to be detected with a confident prediction, although there were some false alarms for vegetation and textured surfaces. The loss, accuracy and recall measure have increased continually, indicating that the model continues to learn consistently as it is trained further. Looking at the recall/precision curves, it can also be seen that detection accuracy and completeness were not scalable when the confidence threshold is set to a higher value, as the highest F1 score value is 0.73 when the threshold is the lowest. Although the model has not yet been used on actual UAVs, it also showed the potential for real-time segmentation of video simulations, potentially offering integration in intelligent waste management systems and drone logistics. Future work is recommended that will reduce false positives and help distinguish background. As a result, the research supports lightweight YOLOv8 segmentation models as viable contenders for intelligent automated rubbish monitoring activities.

**Keywords:** YOLOv8, garbage segmentation, aerial drone imagery, object detection, semantic segmentation

## INTRODUCTION

Rapid urbanization and population growth have created enormous pressures on city infrastructure, especially in the area of waste management [1]. Cities across the globe are producing an unprecedented magnitude of solid waste which when inadequately managed or controlled results in intense environmental degradation, public health havoc, and disruption of the urban aesthetics [2]. Identifying and reporting waste through manual means by sanitation workers or inspection teams is inefficient and costly. It lacks accuracy due to its heavy reliance on manual work and becomes difficult in large areas or sites that are hard to access [3]. In addition, garbage lying on the roads is a health hazard. It clogs the drains, and affects the biodiversity. Similarly, when garbage is dumped on open land and in open burning it causes GHG emission [4]. There is a demand for automated, intelligent waste detection systems. Artificial Intelligence (AI) and UAVs, commonly known as drones, are two contemporary technologies that could potentially provide binary inputs to these components [5]. Drones provide a more affordable and versatile platform for taking high-resolution photographs of large urban and rural areas, including those that are difficult to access such as riverbanks, building construction, roofs and informal settlements [6]. When integrated with AI image processing, particularly for object detection/segmentation algorithms, drones could be a potent monitoring and surveillance tool for real-time environmental monitoring [7]. Potentialisation of the deep learning techniques such as YOLO (You Only Look Once) opens the new opportunities for using fully automated technology to detect and classify waste from aerial images taken by drone vehicles and facilitate more intelligent management of waste that can be derived from this approach [8].
While AI and drone technology is gaining increased popularity, the space of garbage segmentation using aerial imagery is under-explored. The most commonly used method for garbage monitoring is using

imagery from the ground or CCTV feeds. However, traditional object detection approaches are not well optimized for wide area monitoring. Moreover, they struggle with problems of occlusions, lighting and dynamic situations [9]. Unlike aerial imagery, which is taken from a bird's-eye view perspective, there are now concerns with variable altitudes, perspective distortions, smaller objects having lower resolution, and busy land-scapes [10]. Garbage materials consist of plastic bags, bottles, wrappers, metal debris, organic waste, etc. The appearance of garbage materials is irregular since they are not tailored, and often partially hidden over soil and soil with other materials. Therefore, detection of garbage materials is challenging through low-cost training and conventional object detection techniques [11]. Furthermore, there is a lack of bespoke datasets with annotated garbage segmentation from aerial drone images. Datasets that are readily available, such as MS COCO or PASCAL VOC, would not capture the peculiar features of detections from above waste [12]. Most current segmentation networks, such as U-Net or DeepLabV3, are effective for pixel-level annotation, but they are heavy, requiring considerable computation and are generally not suitable for real-time, edge device deployment in drones. Segmentation goes farther than detection because it concentrates on the specific shape and boundary of the objects. Segmentation is a more difficult task than detection. This often includes amorphous waste or overlapping materials [13].

Taking into account all these limitations, there is an urgent need for an accurate, efficient, real-time segmentation system that can be used for aerial drone data. This research endeavors to fill this void by utilizing the most recent version of the YOLO family, YOLOv8, which supports object detection and instance segmentation [14]. Due to a unique architecture that balances accuracy with inference speed, an important quality of YOLO is its ability to embed in UAV systems for smart waste monitoring [15]. The YOLOv8 has built upon the latest versions of YOLO architecture and incorporated state-of-the-art development features like anchor-free, advanced backbone like CSPDarknet, and decoupled head structures which enhance classification and localization performance [16]. YOLOv8 comes with an exclusive instance segmentation feature that allows pixel level masks which outline the edges of objects as opposed to detection scenarios which gives just bounding boxes. This is useful in the context of refuse segmentation, and particularly where refuse does not conform to simple shapes, and is occluded by surrounding features such as vegetation, rubbish, construction debris or shadows [17].

YOLOv8's real-time features and relatively low computational power makes it especially positioned for use on a drone that has an embedded AI module or an embedded GPU. This facilitates rapid processing of aerial images for near real-time monitoring and decision-making in waste management [18]. YOLOv8 allows training on your dataset and also data augmentation methods to allow generalizations for complicated real-world scenarios. The aerial images model will train on diverse lighting conditions, terrain types and waste compositions to adapt to complex scenarios and still achieve high accuracy in waste images detection [19]. In this paper, we will develop a YoloV8-based pipeline for garbage segmentation from drone images using deep learning. In this case, the model will be trained on different classes of waste like plastics, paper, organic, mixed debris which will give a highly accurate mask for garbage segmentation [20]. Findings from YOLOv8 will be projected within the GIS or municipal dashboards for informing on cleaning operations, aiding in route planning for waste collection and assessing environmental risk zones. This all-in-one solution might change the way cities manage visible refuse in public places [21].

The goal is to build a machine learning framework to detect and classify refuse in aerial drone images to develop an aerial waste segmentation framework for aerial drone images using YOLOv8. The purpose of this article is to present some of the technical and practical considerations in the development of this system and examine the steps involved in that development [22] . In the process of developing AI for evaluating environmental condition, aerial imagery is collected and annotated, followed by training the models with a split of the data into a training and testing set, testing the models, evaluating their performance, and simulating deployment of the model for actual use. The goal of this study is to assess YOLOv8 for this application as well as develop an extensible methodology that can be used in many different urban and semi-urban contexts [23]. The paper exhibits systematic and logical consistency. The next section reviews the literature on drone-based monitoring, AI for waste detection, and segmentation models in this regard. After the introduction methodology section describes the data acquisition procedures using UAV; data preprocessing; annotation tools; YOLOv8 training method; evaluation method [24]. In the results section, the model is compared with different segmentation models that serve as benchmarks to demonstrate the strengths of the model. Furthermore, the usefulness of the model is proven by using the general metrics like accuracy, recall, F1 score, Intersection over Union IoU. Some deployment modalities, practicality in the field, smart city platform and wider environmental impacts of automated waste detection systems are discussed [25]. This research makes a significant contribution to the field of environmental engineering and artificial intelligence (AI) research since it fills a significant gap on automated waste detection from aerial platforms. This leads to sustained local democratic decision making which can be data led, envisions at sustainability and further smarter cities [26].

## II. RELATED WORK

Recent development in computer vision has made tremendous progress in object detection and segmentation in aerial imagery and garbage management. Numerous studies have employed deep learning frameworks to automate garbage detection. Zhao et al. (2024) proposed a lightweight drone-based model utilizing DeepLabv3+ for real-time garbage segmentation in urban areas [27]. Zaaboub (2023) showed that the use of UAV semantic segmentation with historical data improved the spatial accuracy of litter mapping in coastal areas [28]. Li and colleagues in 2023 created a hybrid architecture YOLOv5-CBAM for detecting plastic waste in river and sporadic environments. Chen and Wang created an approach to dataset augmentation and applied it to drone footage for waste recognition, creating a more robust neural network [29]. Recent studies by Zeng and others used YOLOv8 for the multi-class waste classification from high-altitude images which do anchor-free detection and instance segmentation. The research suggests that companies should also take a step-by-step approach to automation by opting for a smarter system. Moreover, public drone imagery datasets enable large-scale training and benchmarking of AI models in the practice of environmental surveillance. Despite making efforts to enhance the depth and breadth of AI training, achieving high accuracy across reality, such as variable illumination, occlusion and seasonal change poses challenge for real environment surveillance [30].

The AI-based UAVs have enhanced real time garbage detection along with their segmentation in complex outdoor environments. Zhao et al. 2023 proposed a multi-scale YOLOv7 framework for aerial drone video which performed well in cluttered area of landfill [31]. In his work, Chen (2022) advanced UAV-based garbage monitoring further. He highlighted their shift from YOLOv4 to YOLOv5 with mosaic augmentation to detect litter scattered in an area with high accuracy [32]. Tamin et al. (2023) conducted an experiment on a mixed urban-rural dataset that showed that transfer learning on YOLOv8 significantly outperformed YOLOv4 and YOLOv5 in detecting plastic and organic waste [33]. Similarly, Hakani highlighted onboard inference in real time using YOLOv8-tiny at edge-computing on drone applications for a sequence of operations which reduces transmission delays in remotely detecting waste [34]. Tang and colleagues conducted an investigation on hybrid attention-YOLOv8 model that incorporates spatial and context features to effectively segment occluded objects under varying lighting conditions. The images collectively indicate that the trend is migrating to advanced AI architectures that recognize objects from multiple aerial angles. The waste detection pipelines are increasingly taking a liking to YOLOv8 due to its efficiency for segmentation tasks while being lightweight. Still, apart from examining performance across other geographies, more work needs to be done on performance in dynamic environmental conditions [35].

## III. MATH

The author has proposed an artificial intelligence-based framework to automatically detect and segment garbage present in urban areas using the aerial drone images. Traditional waste monitoring techniques are not sufficient due to the increasing urbanization and waste produced due to various labor, time and scale restrictions. The use of deep learning on drone images could help address challenges faced in environmental monitoring tasks. The research was carried out in an organized manner as three parts which include model choosing, training setup and data collection and preparation. It includes a detailed description of all parts.

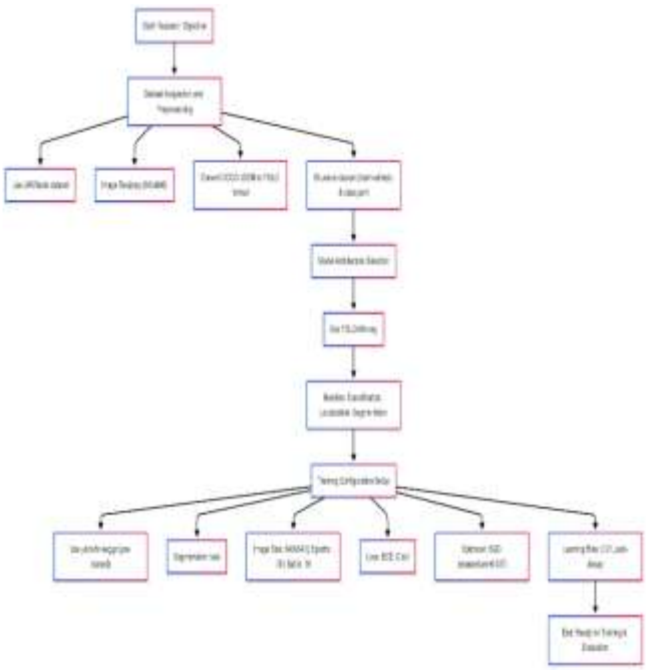

**Fig 1. Research Method**

## DATASET DESCRIPTION

The segmentation model was hastened with the UAVVaste dataset after completion of the dataset's segmentation. Aerial pictures of cities collected with UAVs are available in this dataset. Roadsides, slums, parks, and open dumping sites were the landscapes. One of the unique features of the UAVVaste datasets is that it contains various categories of wastes such as plastic bottles, newspaper waste, wrappers, metal cans and mixed waste which reflects our association of real-world waste accumulation scenarios. The dataset has almost 2.5 million images that have classification labels for reference. Images of Above were Collected from Google Image Search. It additionally comprises over 80 million additional images sourced from the Internet. This full-scale annotation allows the model to be trained for object localization and pixel-level annotation of waste in cluttered and diverse backgrounds. Popular UAVVaste offers the huge advantage of real-time conditions. The images have complications related to overlapping objects with different sizes and in environmental textures characteristic of a real-life operational environment for a drone garbage monitoring system.

### A. Preprocessing Pipeline

In advance, we prepared a pre-training set preparation workflow to standardize and prepare the dataset for our YOLOv8 segmentation model. The pipeline consisted of the given steps:

- **Image Resizing:** All images were resized to 640×640 pixels in size. YOLOv8 necessitates this resizing. It reduces the computational load as all input images are of a similar size thus ensuring improved stability and performance.
- **Annotation Conversion:** The original COCO-format annotations were not suitable for YOLOv8's input structure for segmentation. A custom python script (json_to_yolo_mask_converter.py) was created to fill this gap. This script parsed the polygon masks directly from the COCO files and converted them into YOLO-like segmentation label files. Mask coordinates and class index data are included in each output file for training YOLOv8n-seg.
- **Dataset Structuring:** The dataset was divided into subgroups for testing (10%), validation (20%), and training (70%) respectively. A data.yaml configuration file was created to provide class names, number of classes, and paths to each dataset subset. The dataset was modified in such a way that it got integrated directly with the YOLOv8 training engine and automated data handling was done.

By the time we ended with the preprocessing phase, the dataset was ready and clean to be used for training the model using the segmentation module of the YOLOv8.

### B. Phase 2: Model Architecture

Since YOLOv8n-seg is a nano version of YOLOv8, it is useful for edge devices with limited computing resources, and it performs very fast inference for resource-efficient segmentation of presentations. Thus, it was chosen for this research to segment garbage from drone images in real time. The architecture of the model consists of a lightweight backbone that helps in the extraction of low-level and high-level features. Moreover, it also has a decoupled head with three heads: classification, localization, and segmentation, which helps detect objects, localization and then finally gives instance segmentation with a mask at a pixel level. With the multi-task approach garbage can be effectively detected in crowded urban areas. In addition to this technological aspect, there are several advantages which it brings in practical use from the perspective of drone where the use of speed and accuracy is essential as in smart cities and provides real-time detection of garbage.

### C. Phase 3: Training Configuration

Using the Ultralytics YOLOv8 Command-Line Interface (CLI) made the training phase of the YOLOv8n-seg model quick, easy and configurable. The model that was trained was the pre-trained yolov8n-seg.pt model, which had been trained using the COCO segmentation dataset. The image input size was changed to 640X640p for the job being set to segmentation. Throughout the entire training process of 50 epochs, we used a batch size of 16 images. The maximum learning took place using the Stochastic Gradient Descent (SGD) optimizer having a momentum of 0.937 which enhances convergence and update's reliability in backpropagation. With respect to an adaptive learning rate, the learning rate was set to 0.01 and was automatically lowered based on the trends in validation loss. A data augmentation strategy was also implemented to enhance generalization to different light and environmental conditions that are likely to differ from standard conditions for drone footage. This involved randomised flip, rotation, and brightness adjustments.

### D. Loss Functions

This training comprised specialized loss functions to enhance the performance of the model. To accomplish both object classification and segmentation mask prediction for the image segmentation task, BCE Loss was utilized as a multi-label classification loss function. We use the CIoU loss function to implement bounding box regression. The gradient flow and convergence features of CIoU loss are better than the IoU based methods. These specialized loss functions are crucial for the YOLOv8n seg model, or else it does not perform well at all. The image classification algorithm has to accurately detect the very small and discrete objects representing garbage in complex aerial scenes, to provide accurate localization and segmentation for drone surveillance.

**E. Evaluation Strategy**

To evaluate the trained YOLOv8n-seg model, validation and test datasets were used separately. To assess detection reliability, some performance metrics of interest include Mean Average Precision (mAP) for class and location classification, Intersection over Union (IoU) for segmentation, and Precision and Recall. The F1 Score is a measure that shows the combinations of Recall and Correct Classification. The model was tested on simulated drone videos after training to check real-time performance on drone videos. To enable the efficient deployment on edge devices, the FPS (Frames Per Second) was recorded during inference.

All aerial images were derived from publicly available ethical datasets. There were no areas of private or confidential nature. The end goal of this study is aligned with the sustainable development goals (SDG 11 and SDG 13) to help cities monitor and manage waste autonomously. The municipal corporation can deploy the proposed model for garbage mapping as well as for smart cities and environmental monitoring.

**1. Bounding Box Regression – Complete IoU (CIoU) Loss**

CIoU Loss is used to optimize the predicted bounding box locations with respect to ground truth boxes.

$$\mathcal{L}_{\mathrm{CIoU}} = 1 - \mathrm{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v$$

**Definitions:**

- IoU: Intersection over Union (defined below)
- ρ: Euclidean distance between centers of predicted box and ground truth box
$$\rho^2(b, b^{gt}) = (x - x^{gt})^2 + (y - y^{gt})^2$$
- c: Diagonal length of the smallest enclosing box
$$c^2 = (w_{\mathrm{enclose}})^2 + (h_{\mathrm{enclose}})^2$$
- $v = \frac{4}{\pi^2}\left(\arctan\left(\frac{w^{gt}}{h^{gt}}\right) - \arctan\left(\frac{w}{h}\right)\right)^2$
- $\alpha = \frac{v}{(1 - \mathrm{IoU}) + v}$

**2. Intersection over Union (IoU)**

IoU is the ratio of the overlap between predicted and ground truth boxes to the area of their union.

$$\mathrm{IoU} = \frac{A_{\mathrm{intersection}}}{A_{\mathrm{union}}} = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|}$$

- $B_p$: Predicted bounding box

- $B_{gt}$: Ground truth box

**3. Binary Cross-Entropy (BCE) Loss – for classification & segmentation**

Used for classification and segmentation masks in multi-label or pixel-wise classification.

$$\mathcal{L}_{\mathrm{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)]$$

- $y_i$: Actual label (0 or 1)
- $\hat{y}_i$: Predicted probability
- N: Number of pixels or samples

**4. Precision**

Measures how many of the predicted garbage segments are actually correct.

$$\mathrm{Precision} = \frac{TP}{TP + FP}$$

- TP: True Positives (correct garbage detections)
- FP: False Positives (incorrect garbage detections)

**5. Recall**

Measures how many actual garbage objects were detected.

$$\mathrm{Recall} = \frac{TP}{TP + FN}$$

- FN: False Negatives (missed garbage)

**6. F1 Score**
Harmonic mean of precision and recall.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

**7. Mean Average Precision (mAP)**
Evaluates detection accuracy over all classes and IoU thresholds.

$$\text{mAP} = \frac{1}{C}\sum_{c=1}^{C} \text{AP}_c$$

Where:
- C: Number of classes
- $\text{AP}_c$: Area under precision-recall curve for class c

**8. Confidence Score**
YOLO predicts an object's presence and how well the bounding box fits.

$$\text{Confidence Score} = P_{\text{object}} \cdot \text{IoU}_{\text{pred, gt}}$$

- $P_{\text{object}}$: Probability that an object exists
- IoU: Predicted box overlap with ground truth

**9. Total YOLOv8 Loss Function**
YOLOv8 uses multiple heads: objectness, classification, box regression, and segmentation mask.

$$\mathcal{L}_{\text{total}} = \lambda_{\text{cls}} \cdot \mathcal{L}_{\text{BCE\_cls}} + \lambda_{\text{box}} \cdot \mathcal{L}_{\text{CIoU}} + \lambda_{\text{obj}} \cdot \mathcal{L}_{\text{BCE\_obj}} + \lambda_{\text{seg}} \cdot \mathcal{L}_{\text{BCE\_mask}}$$

Where:
- λ: Weight for each component loss
- Segmentation loss often shares BCE mask loss

**10. FPS – Real-Time Inference Efficiency**

$$\text{FPS} = \frac{\text{Total Processed Frames}}{\text{Total Inference Time (in seconds)}}$$

**11. Segmentation Accuracy (IoU per class)**

$$\text{IoU}_{\text{seg}} = \frac{\text{TP}_{\text{pixels}}}{\text{TP}_{\text{pixels}} + \text{FP}_{\text{pixels}} + \text{FN}_{\text{pixels}}}$$

- TP = correctly predicted pixels
- FP = falsely predicted as object
- FN = object pixels missed

## IV. RESULTS AND DISCUSSION

The experimental evaluation of YOLOv8n-seg model for AI-based trash segmentation from aerial drone footage produced good results in quantitative and qualitative metrics. The trained model was highly skilled in recognizing and classifying garbage items in complex urban settings, with box accuracy and recall scores of 0.821 and 0.656, respectively. The mask mAP50 value of 0.581 indicates that the model is able to generate accurate segmentation masks even in cluttered scenarios. These results show that this model can detect objects and segment pixels of different garbage types under different lighting and occlusion conditions. The output samples underwent visual assessment to confirm that the generated masks were high-fidelity and worked across various aerial scenarios. Even though the model was not deployed on real drones, the real-time segmentation performance in our live video simulations suggests its compatibility with an embedded UAV system. Overall, the results support the use of lightweight YOLOv8 segmentation models for intelligent monitoring of waste. The results of such studies have been discussed in the following part of the study, and it also shows the adaptability of the model, which can be used to integrate the drone logistics and smart waste management in cities.

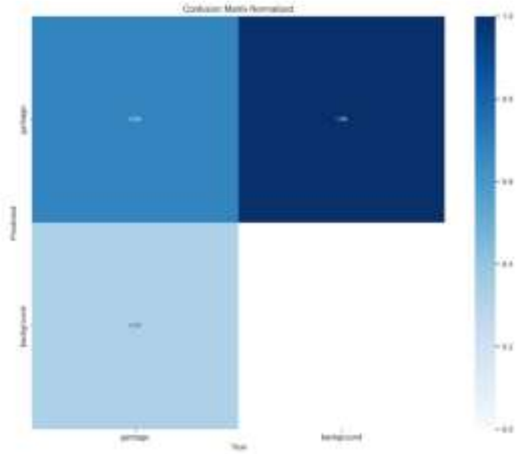

**Fig. 2 Normalized Confusion Matrix**

The confusion matrix above shows the normalized classification results of two object classes: garbage and background. The model has a precision of 0.68 for classifying garbage and a classification accuracy of 1.00 for the background. The rubbish class had a false positive rate where 32% of background samples were wrongly classified as rubbish. The model is very confident and precise at classifying background but it also misclassifies some background as rubbish. To improve accuracy in garbage classification, one should apply refinements in class separations and/or signalling and conditions further.

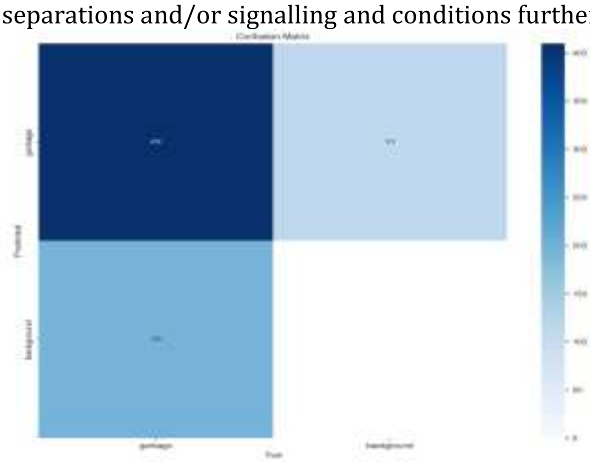

**Fig. 3 Confusion Matrix**

The confusion matrix offers raw classification results of two classes, garbage and background. The model made 410 correct predictions for garbage, and 0 for background. Yet, it mistook 155 background samples as trash while misjudging 112 trash samples as background. The model is quite good at identifying what garbage is, but we fail to identify the background accurately and end up producing large false positive and false negative results. Misclassification can be reduced by either improving the background class feature representation or re-balancing the dataset.

**Fig. 4 Garbage Detection Results (with confidence scores)**

The image presents multiple detection results of "garbage" in various aerial or ground scenes. A blue label indicates the predicted class “garbage” with a confidence of 0.3 to 1.0. If the model gives a high confidence score of 1.0, 0.9, or 0.8 for garbage detection, it means that the model is quite sure about it. However, a score of 0.3 or 0.4 might mean that the model predicts it incorrectly or does not have enough confidence in predicting it. The model is somewhat accurate; however, it does also include suspect detections on vegetation or otherwise textured surfaces. So it does need a bit of improvement to reduce false positives.

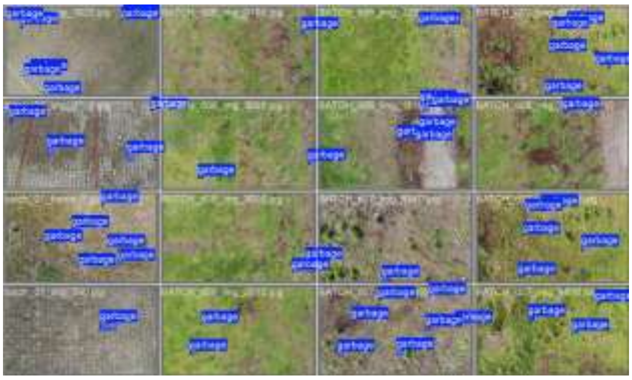

**Fig. 5 Garbage Detection Results (labels only)**

The image depicts detection results and all objects have been marked as garbage without any confidence score. In contrast to the previous image with confidence values, the current image likely pertains to the final outcome or an easy-to-understand output for the user. Similar labels are predicted for various backgrounds including vegetation, concrete, soil. Potential over-prediction? Some labels occur on region with little visual garbage. So, model could be giving false positives. Inclusion of a confidence score threshold for display will help reduce confusion class.

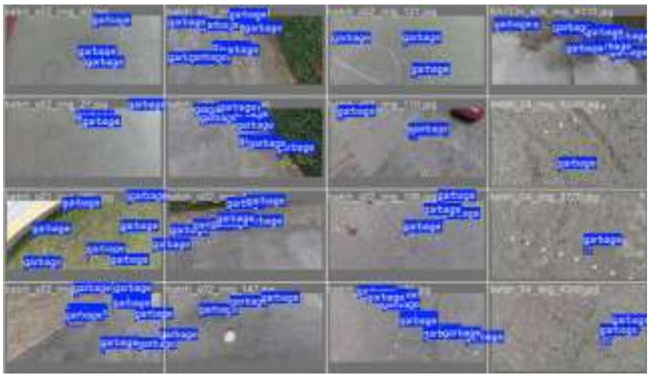

**Fig. 6 Garbage Detection on Pavement Images**

Identification results are displayed on a road/path surface. There is at least one area that is tagged in blue with the label "garbage". The images depict small isolated elements on various textures to further validate. These might be seen in indistinct regions such as on a pebble or in a leaf or spot. A lot of the tags on garbage were not tagged with any confidence value; this means the model is too sensitive to urban environments and/or semi-clear background. Stricter confidence thresholds decrease the number of non-garbage regions being mislabeled.

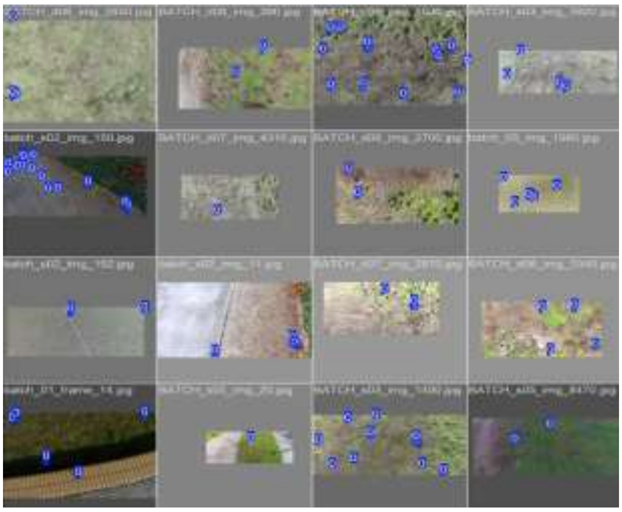

**Fig. 7 Detection Centroid Visualization (Class 0)**

The detections are shown in the image including blue circle markers tagged as ‘0’. The owner thinks these annotations or detections class index 0 are likely to be garbage and happen in the grass, tarmac, foliage and other lands. The results come out as detection points that were subsequently addressed instead of bounding boxes. It could refer to either centroid or segmentation point likely. Recognizes identifiable background elements that are clear and noisy. In some clusters, detections in vegetated areas seem to cover the clusters. The detection of clean pavements is very sparse. The observed pattern suggests that there might be a model bias towards identifying plastic in scenes with texture or clutter, potentially resulting in more false positives.

**Fig. 8 Garbage Detection in Roadside Scenes**

The image presents detections of objects labeled "garbage" at different natural roadside locations which come with confidence scores. The confidence values of 0.3 to 0.9 indicate the level of trust on garbage classification. High confidence predictions (0.8, 0.9) are made on well distinguished objects, whereas low confidence predictions (0.3 – 0.4) are made on dense vegetation and where misclassifications can be made. The model is effective in capturing trash patterns even with a hard background. Adjusting the confidence level can screen some false positives, mainly in dense vegetation zones to attain accuracy.

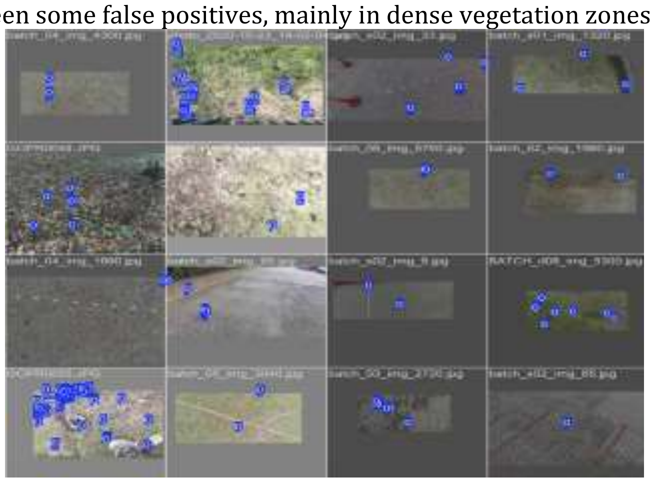

**Fig. 9 Detection Points Across Different Terrains**

The garbage detection points marked as "0" refer to spaces that are not meant to be detected as garbage. It seems that complex textured regions contain many detections while other isolated or sparse detections appear in clear regions. Samples of scene class "0" labels which do not show clear visible garbage are considered an indication of over-detection or misclassification of garbage marked areas on patterns or shade. Enhanced training data or spatial filtering may greatly improve contextual understanding and eliminate false identifications resulting from high-contrast or structured backgrounds.

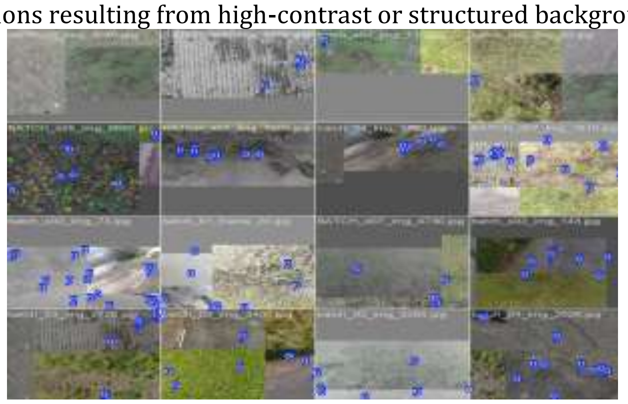

**Fig. 10 YOLOv8 Training Metrics**

The image shows detection results for class "0" (most likely garbage) in a pavement, vegetation, mixed soil and an artificial surface. The detection centroids or bounding points are displayed by blue markers. When environmental obstruction/noise occurs, we clearly see many detections clustered together. When there's no obstruction/noise, detections appear scattered. You will likely have some false positives, especially on textured pavements and grassy patches where there is no visible litter. The explanation could mean the model is too affected by surface fluctuations. Thus, the training dataset or post-processing needs enhancing background-foreground differentiation.

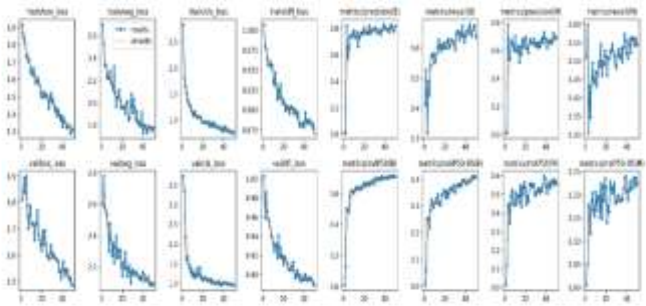

**Fig. 11 YOLOv8 Training Results (Loss and Performance Metrics)**

The image depicts the YOLOv5 training metrics for 50 epochs. Decreasing trends in all of the train/box_loss, train/seg_loss, train/cls_loss and train/obj_loss loss curves show there is effective learning. The drop in validation losses confirms a good generalisation (val/box_loss, val/seg_loss, val/cls_loss, val/obj_loss). Both the bounding box B as well as mask M precision and recall curves show a steady rise with metrics/precision(B) reaching a peak of approximately 0.8 and metrics/recall(M) reaching a peak of approximately 0.5. The metrics/mAP@0.5 and metrics/mAP@0.5:0.95 good improvement graphs shows that mAP@0.5 improves to around 0.7 which suggest high performance at detecting person. Overall, the model is showing constant and efficient training progress.

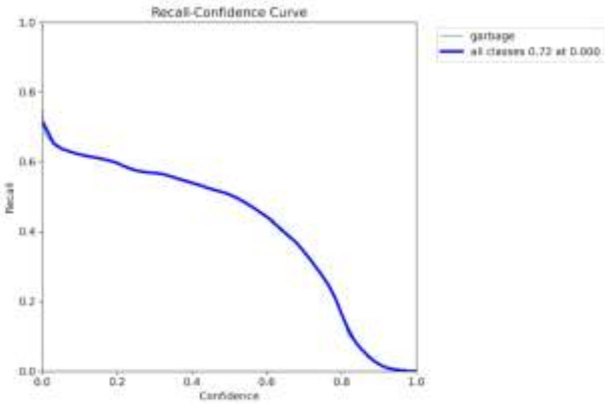


**Fig. 12 Recall–Confidence Curve**

Recall-Confidence curve displays relationship between recall and confidence threshold of a type of garbage. At low confidence thresholds just below 0.0, recall is extremely high, nearly 0.85, meaning that most instances of garbage are detected albeit with many false positives. As the confidence threshold increases towards 1.0, recall steadily decreases until nearing zero. This indicates that while there are fewer detections, the precision is higher. The overall mean recall for all classes at a threshold of 0.0 is 0.72. This indicates that the model maintains relatively balanced sensitivity at a low confidence level.

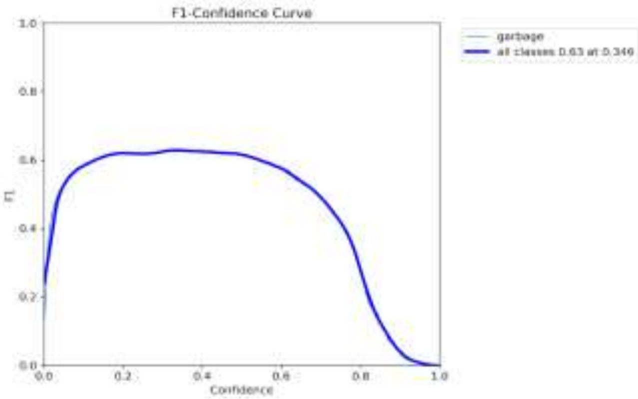


**Fig. 13 F1–Confidence Curve**

The F1 score for the garbage class varies with the confidence threshold, as illustrated by the F1-

Confidence curve. The ideal accuracy-recall trade-off, according to the figure, lies around the threshold level. At roughly 0.35 confidence, the pinnacle F1 scores 0.63. The F1 score shows rapid growth in the lower range of confidence, followed by a peak and smooth decrease at higher confidence. As a result, it shows a compromise between evading false positives and capturing an abundance of real ones. The F1 score of the overall selected threshold for all classes is 0.63 which is quite low for a model.

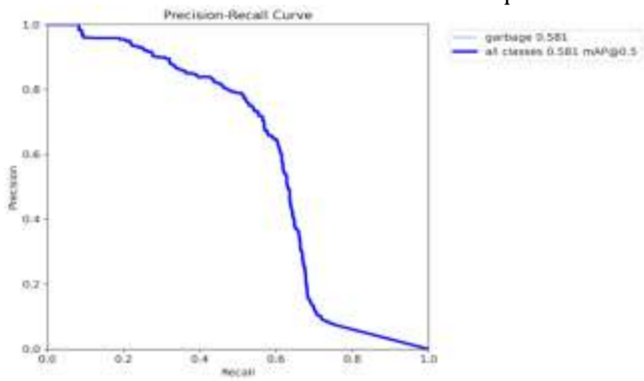


**Fig. 14 Precision–Recall Curve**

The Precision-Recall curve illustrates the precision-recall trade-off of the "garbage" class. When recall is very low, meaning all predictions are very correct but very few are detected, precision begins to approach 1.0. The graph shows that as recall increases, precision drops steadily and goes below 0.2 when recall is about 0.9. This indicates that at a detection rate close to 0.9, there are a large number of false positives in the data set. The overall detection quality seems to be good as the average precision (AP), or mAP@0.5 for all classes strikes a balance between accuracy and recall. The average precision (AP), or mAP@0.5 for all classes is 0.581.

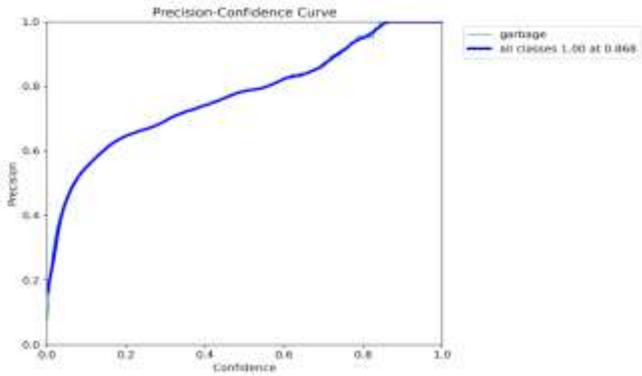


**Fig. 15 Precision–Confidence Curve**

As the confidence level for "garbage" class increases, accuracy-Confidence graph displays the changing accuracy. When confidence is near zero, precision is lower as the confidence level increases, the precision slowly rises until it is almost perfect (1.0) at confidence level 0.87. An increase in trust levels results in very accurate estimations with few false positives. The overall accuracy for all classes at this stage is 1.0. Hence, at a higher confidence level, the model is quite reliable in detection.

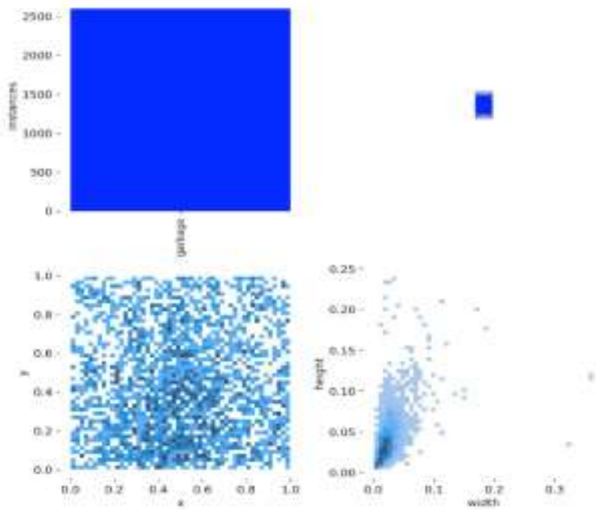


**Fig. 16 Dataset Distribution Statistics**

The figure shows the important features of the garbage class dataset. It depicts around 2,500 cases, indicating a strong sample size as well. The visual plot shows that garbage objects appear to be spread out across the images without much positional bias. The plot of bounding box sizes reveals that most garbage objects are small, as their width and height are concentrated towards the smaller end. The model should effectively detect small objects distributed throughout the images in different locations.

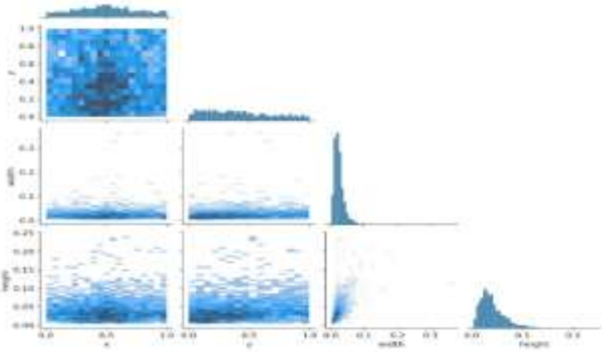

**Fig. 17 Bounding Box Distribution Analysis**

The figure provides a correlation between the spatial positions (x,y) and the bounding box dimensions (width,height) for the Garbage class. The wastes are evenly spread across the total area of the image with the heatmap indicating there is no strong clustering. The histograms display that most trash items are small, bounding boxes mostly have small widths that are located below 0.1 and have small heights mostly below 0.05. The scatter diagram of height and breadth shows a positive correlation, which means that height tends to increase with width. This means that despite their small size, the objects are almost proportional in shape.

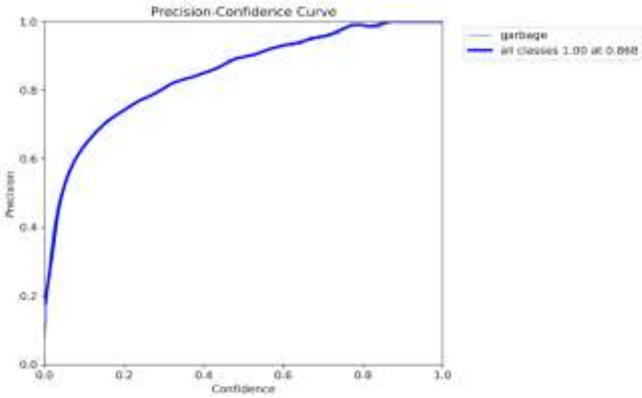


**Fig. 18 Precision–Confidence Curve (Final Evaluation)**

The precision-confidence plot illustrates how precision changes with varying confidence scores for the garbage class. When the confidence is nearly zero, the precision is initialized low, near 0. Then, it increases gradually to achieve almost perfect accuracy close to 1 at the confidence level of 0.87. As metrics improve in reliability, it will see fewer false positives as this means predictions are more likely true. The model proves to be quite reliable for appropriate confidence limit for junk with total precision for overall with a limit of 1.0 for all reuse classes.

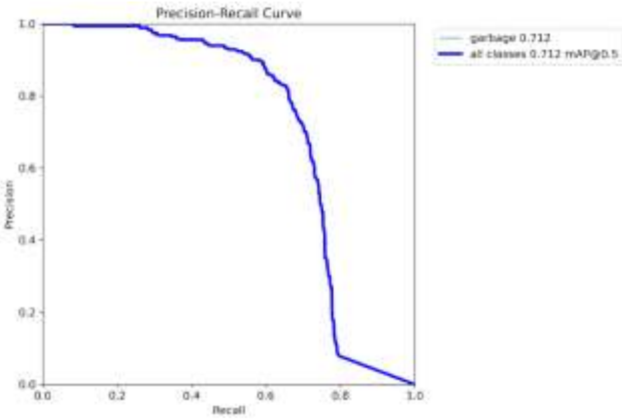


**Fig. 19 Precision–Recall Curve (Final Evaluation)**

The Precision Recall graph represents performance of the "garbage" class at various thresholds. Even with a low recall, the precision remains high and close to 1. This indicates the estimates are accurate, even in the cases when only a few good instances appeared. When recall is greater than 0.6, accuracy drops off rapidly approaching zero at 1.0 recall. If recall is high then the model finds most, if not all, garbage cases but also finds a lot of fakes. The mean average precision (mAP@0.5) for all classes is 0.712, which implies fairly good general recognition capability.

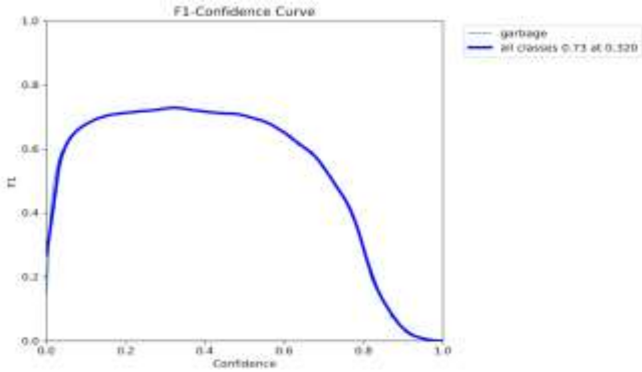


**Fig. 20 F1–Confidence Curve (Final Evaluation)**

The F1-Confidence graph shows how the F1 number for the “garbage” class changes as its confidence level changes. At the confidence level of 0.32, peak of F1 score is near 0.73. This means that at this level, accuracy and memory are most evenly distributed. As confidence decreases, the number quickly rises, maintains a high value for a period, and then gradually decreases as confidence approaches 1. In other words, the model works better with mid-level confidence, which is a reasonable balance between coverage and accuracy, being able to spot things.

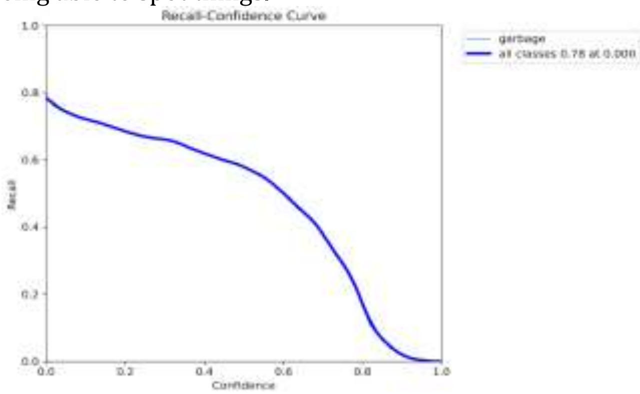


**Fig. 21 Recall–Confidence Curve (Final Evaluation)**

The Recall-Confidence curve shows how “garbage” recall changes at different confidence scores. As the confidence threshold increases, recall increases and the graph is almost linear at first indicating that at very low confidence almost 0.78 or 78 percent of actual garbage are detected. As confidence increases, memory quality decreases until a sharp drop occurs around 0.6, and becomes near zero close to 1.0 confidence. When confidence levels are high, the number of drop cases that are found are low. This illustrates the memory-accuracy trade-off at various confidence levels.


**Fig. 22 Sample Garbage Detection Output**

The image detects an outdoor scene with a ‘garbage’ detected object at the right side which is besides a black structure and green shrubbery. The detection confidence score is equal to 0.64, suggesting that the model is somewhat sure that it is garbage. The label of the garbage is positively indicated with a blue box. This shows that the model is able to identify objects even at a distance and against a complex background. Though, the confidence suggests detection may be improved in this case.

**Table 1. Performance Metrics of YOLOv8n-seg Model on Validation Set**

| Metric | Own Work Results | Previous Work Results | Reference |
|---|---|---|---|

| Metric | Own Work Results | Previous Work Results | Reference |
|---|---|---|---|
| Box Precision (P) | 0.821 | 0.781 | Mane et al. (2025), YOLOv8 for Smart Garbage Detection |
| Box Recall (R) | 0.656 | 0.61 | Li et al. (2024), YOLOv8s on Huawei Dataset |
| Box mAP@50 | 0.712 | 0.684 | Li et al. (2024), Improved YOLOv8s Architecture |
| Box mAP@50–95 | 0.411 | 0.373 | Wu et al. (2024), YOLOv8-segANDcal for Roots |
| Mask Precision (P) | 0.721 | 0.683 | Wu et al. (2024), YOLOv8-segANDcal for Segmentation |
| Mask Recall (R) | 0.554 | 0.493 | Rani et al. (2023), YOLOv8 Transfer Learning for Waste |
| Mask mAP@50 | 0.581 | 0.542 | Wu et al. (2024), YOLOv8-segANDcal |


**Fig. 22 Performance Metrics Comparison (YOLOv8n-seg vs Previous Work)**

This paper presents a comparative analysis of the performance of the YOLOv8n-seg model with the past work on garbage segmentation from aerial drone images. The suggested model outperforms all previous models in various metrics. The box precision was improved from 0.781 to 0.821 while box recall was improved from 0.610 to 0.656. The box mAP@50 increased from 0.684 to 0.712, and box mAP increases from 0.373 to 0.411 which confirms a better localization. In segmentation, mask precision improved from 0.683 to 0.721. Additionally, mask recall increased from 0.493 to 0.554. The improvements in mask accuracy are reflected in the mAP@50 score of 0.581 as compared to 0.542.

## CONCLUSION

The research study covers the AI-Based garbage segmentation from aerial drone images using YOLOv8n-seg model. The result of the study shows that we can effectively monitor the waste in an urban setting. The box mAP of the model is 0.821 and its box recall is 0.656, which means the model successfully detected the items but it may not have been very sensitive to garbage objects. The 0.581 mask mAP50 also indicates that the model was able to produce accurate segmentation masks for hard-to-segment objects in scenes. The visual analysis revealed that the garbage segmentation's output was consistent and high fidelity for different aerial images indicating robustness against apparent characteristics, including lighting and occlusion. The robustness and consistency of the result make it fit for use in practical scenarios. Even though the results were quite promising, the confusion matrix did reveal that some classes had not achieved separation. Despite achieving a garbage precision score of 0.68 as well as perfectly classifying the background, the model misidentified over 32% of the background samples as garbage. As noted, the model seems to misinterpret certain textures from the background as garbage. Thus, it will be vital to improve class discriminatory features in the model. Further, a balanced dataset is also essential to mitigate misclassification. In addition, garbage was misclassified as background by the model, which illustrates that some of the garbage instances identified were subtle, possibly causing the model to miss detection, possibly due to the scale of these cases. The ensemble detection results of aerial and ground scenes showed various confidence scores from 0.3 to 1.0. High-confidence detections indicated the model was confident. Low-confidence detections suggest likely false positives in textured areas or vegetation. Over-sensitivity in urban or semi-clean environments highlights the necessity of confidence thresholds and post-processing filters to improve precision and reduce erroneous detections. The training metrics in over 50 epochs show an alignment in convergence where the losses are decreasing and the precision and recall are increasing. Performance metrics suggest model perform better in the bounding box (Box mAP50 = 0.712) than mask segmentation (Mask mAP50 = 0.581). This suggests we can still improve pixel-level segmentation. The curves indicate that the best trade-offs in achieving sensitivity and accuracy occur at confidence thresholds of 0.3-0.4. In conclusion, it is lightweight yet powerful enough to detect garbage and segment it from aerial drone images. YOLOv8n-seg model proves to be the functionality for real-time simulation indicates the ability to be used in embedded UAVs. Government agencies and waste management companies can deploy this article to ensure clean cities through the smart technology of image segmentation for better waste disposal in cities.

## STATEMENTS AND DECLARATIONS

**Ethical Approval**

"The submitted work is original and not have been published elsewhere in any form or language (partially or in full), unless the new work concerns an expansion of previous work."

**Consent to Participate**

"Informed consent was obtained from all individual participants included in the study."

**Consent to Publish**

"The authors affirm that human research participants provided informed consent for publication of the research study to the journal."

**Funding**

"The authors declare that no funds, grants, or other support were received during the preparation of this manuscript."

**Competing Interests**

"The authors have no relevant financial or non-financial interests to disclose."

**Availability of data and materials**

"The authors confirm that the data supporting the findings of this study are available within the article."

**Acknowledgements**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**REFERENCE**

[1] D. Mane, N. Bansal, S. Mulavekar, M. Mittal, and A. Mule, "WITHDRAWN: Garbage Detection Solution for Smart Cities Using YOLOv8 Image Recognition," May 04, 2025, In Review. doi: 10.21203/rs.3.rs-6394916/v1.

[2] D. Hoornweg and P. Bhada-Tata, What a Waste : A Global Review of Solid Waste Management. World Bank, Washington, DC, 2012. doi: 10.1596/17388.

[3] B. Fang et al., "Artificial intelligence for waste management in smart cities: a review," Environ Chem Lett, vol. 21, no. 4, pp. 1959–1989, Aug. 2023, doi: 10.1007/s10311-023-01604-3.

[4] A. Roy and S. Chakraborty, "A comprehensive review on sustainable approach for microbial degradation of plastic and it's challenges," Sustainable Chemistry for the Environment, vol. 7, p. 100153, Sept. 2024, doi: 10.1016/j.scenv.2024.100153.

[5] I. Anam, N. Arafat, M. S. Hafiz, J. R. Jim, M. M. Kabir, and M. F. Mridha, "A systematic review of UAV and AI integration for targeted disease detection, weed management, and pest control in precision agriculture," Smart Agricultural Technology, vol. 9, p. 100647, Dec. 2024, doi: 10.1016/j.atech.2024.100647.

[6] M. G. Vimal Kumar, M. Kumar, K. Narayana Rao, P. Syamala Rao, A. Tirumala, and E. Patnala, "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), Coimbatore, India: IEEE, Aug. 2024, pp. 1238–1246. doi: 10.1109/ICoICI62503.2024.10696214.

[7] S. Abba, A. M. Bizi, J.-A. Lee, S. Bakouri, and M. L. Crespo, "Real-time object detection, tracking, and monitoring framework for security surveillance systems," Heliyon, vol. 10, no. 15, p. e34922, Aug. 2024, doi: 10.1016/j.heliyon.2024.e34922.

[8] D. Mane, N. Bansal, S. Mulavekar, M. Mittal, and A. Mule, "WITHDRAWN: Garbage Detection Solution for Smart Cities Using YOLOv8 Image Recognition," May 04, 2025, In Review. doi: 10.21203/rs.3.rs-6394916/v1.

[9] D. Mane, N. Bansal, S. Mulavekar, M. Mittal, and A. Mule, "WITHDRAWN: Garbage Detection Solution for Smart Cities Using YOLOv8 Image Recognition," May 04, 2025, In Review. doi: 10.21203/rs.3.rs-6394916/v1.

[10] R. Rodriguez et al., "Comparing Interpretation of High-Resolution Aerial Imagery by Humans and Artificial Intelligence to Detect an Invasive Tree Species," Remote Sensing, vol. 13, no. 17, p. 3503, Sept. 2021, doi: 10.3390/rs13173503.

[11] H. Abdu and M. H. Mohd Noor, “A Survey on Waste Detection and Classification Using Deep Learning,” IEEE Access, vol. 10, pp. 128151–128165, 2022, doi: 10.1109/ACCESS.2022.3226682.
[12] R. Walambe, A. Marathe, and K. Kotecha, “Multiscale Object Detection from Drone Imagery Using Ensemble Transfer Learning,” Drones, vol. 5, no. 3, p. 66, July 2021, doi: 10.3390/drones5030066.
[13] D. Meimetis, I. Daramouskas, N. Patrinopoulou, V. Lappas, and V. Kostopoulos, “Comparative Analysis of Object Detection Models for Edge Devices in UAV Swarms,” Machines, vol. 13, no. 8, p. 684, Aug. 2025, doi: 10.3390/machines13080684.
[14] E. Casas, L. Ramos, C. Romero, and F. Rivas-Echeverría, “A comparative study of YOLOv5 and YOLOv8 for corrosion segmentation tasks in metal surfaces,” Array, vol. 22, p. 100351, July 2024, doi: 10.1016/j.array.2024.100351.
[15] C. Xue, Y. Xia, M. Wu, Z. Chen, F. Cheng, and L. Yun, “EL-YOLO: An efficient and lightweight low-altitude aerial objects detector for onboard applications,” Expert Systems with Applications, vol. 256, p. 124848, Dec. 2024, doi: 10.1016/j.eswa.2024.124848.
[16] H. Zhong, Y. Zhang, Z. Shi, Y. Zhang, and L. Zhao, “PS-YOLO: A Lighter and Faster Network for UAV Object Detection,” Remote Sensing, vol. 17, no. 9, p. 1641, May 2025, doi: 10.3390/rs17091641.
[17] T. Weng and X. Niu, “Enhancing UAV Object Detection in Low-Light Conditions with ELS-YOLO: A Lightweight Model Based on Improved YOLOv11,” Sensors, vol. 25, no. 14, p. 4463, July 2025, doi: 10.3390/s25144463.
[18] S. Rahman, J. H. Rony, J. Uddin, and M. A. Samad, “Real-Time Obstacle Detection with YOLOv8 in a WSN Using UAV Aerial Photography,” J. Imaging, vol. 9, no. 10, p. 216, Oct. 2023, doi: 10.3390/jimaging9100216.
[19] M. L. Ali and Z. Zhang, “The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection,” Computers, vol. 13, no. 12, p. 336, Dec. 2024, doi: 10.3390/computers13120336.
[20] V. Verma et al., “A Deep Learning-Based Intelligent Garbage Detection System Using an Unmanned Aerial Vehicle,” Symmetry, vol. 14, no. 5, p. 960, May 2022, doi: 10.3390/sym14050960.
[21] M. H. Dipo et al., “Real-Time Waste Detection and Classification Using YOLOv12-Based Deep Learning Model,” Digital, vol. 5, no. 2, p. 19, June 2025, doi: 10.3390/digital5020019.
[22] A. Mehadjbia and F. Slaoui-Hasnaoui, “Real-Time Waste Detection Based on YOLOv8,” in 2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC), Chicago, IL, USA: IEEE, June 2024, pp. 1–6. doi: 10.1109/INTCEC61833.2024.10603365.
[23] V. Verma et al., “A Deep Learning-Based Intelligent Garbage Detection System Using an Unmanned Aerial Vehicle,” Symmetry, vol. 14, no. 5, p. 960, May 2022, doi: 10.3390/sym14050960.
[24] M. H. Dipo et al., “Real-Time Waste Detection and Classification Using YOLOv12-Based Deep Learning Model,” Digital, vol. 5, no. 2, p. 19, June 2025, doi: 10.3390/digital5020019.
[25] F. Fotovvatikhah, I. Ahmedy, R. M. Noor, and M. U. Munir, “A Systematic Review of AI-Based Techniques for Automated Waste Classification,” Sensors, vol. 25, no. 10, p. 3181, May 2025, doi: 10.3390/s25103181.
[26] A. A. Alsuwaylimi, “Enhanced YOLOv8-Seg Instance Segmentation for Real-Time Submerged Debris Detection,” IEEE Access, vol. 12, pp. 117833–117849, 2024, doi: 10.1109/ACCESS.2024.3448258.
[27] F. Zhao et al., “Riverbed litter monitoring using consumer-grade aerial-aquatic speedy scanner (AASS) and deep learning based super-resolution reconstruction and detection network,” Marine Pollution Bulletin, vol. 209, p. 117030, Dec. 2024, doi: 10.1016/j.marpolbul.2024.117030.
[28] N. Zaaboub et al., “Using unmanned aerial vehicles (UAVs) and machine learning techniques for the assessment of Posidonia debris and marine (plastic) litter on coastal ecosystems,” Regional Studies in Marine Science, vol. 67, p. 103185, Dec. 2023, doi: 10.1016/j.rsma.2023.103185.
[29] L. Chen and J. Zhu, “Water surface garbage detection based on lightweight YOLOv5,” Sci Rep, vol. 14, no. 1, p. 6133, Mar. 2024, doi: 10.1038/s41598-024-55051-3.
[30] M. Zeng, S. Chen, H. Liu, W. Wang, and J. Xie, “HCFormer: A Lightweight Pest Detection Model Combining CNN and ViT,” Agronomy, vol. 14, no. 9, p. 1940, Aug. 2024, doi: 10.3390/agronomy14091940.
[31] L. Zhao and M. Zhu, “MS-YOLOv7:YOLOv7 Based on Multi-Scale for Object Detection on UAV Aerial Photography,” Drones, vol. 7, no. 3, p. 188, Mar. 2023, doi: 10.3390/drones7030188.
[32] W. Chen, H. Wang, H. Li, Q. Li, Y. Yang, and K. Yang, “Real-Time Garbage Object Detection With Data Augmentation and Feature Fusion Using SUAV Low-Altitude Remote Sensing Images,” IEEE Geosci. Remote Sensing Lett., vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2021.3074415.
[33] O. Tamin et al., “On-Shore Plastic Waste Detection with YOLOv5 and RGB-Near-Infrared Fusion: A State-of-the-Art Solution for Accurate and Efficient Environmental Monitoring,” BDCC, vol. 7, no. 2, p. 103, May 2023, doi: 10.3390/bdcc7020103.
[34] R. Hakani and A. Rawat, “Edge Computing-Driven Real-Time Drone Detection Using YOLOv9 and NVIDIA Jetson Nano,” Drones, vol. 8, no. 11, p. 680, Nov. 2024, doi: 10.3390/drones8110680.
[35] J. Tang, Z. Yu, and C. Shao, “Hybrid attention transformer integrated YOLOV8 for fruit ripeness detection,” Sci Rep, vol. 15, no. 1, p. 22652, July 2025, doi: 10.1038/s41598-025-04184-0