FMDB Transactions on Sustainable Computing Systems

AI-Powered Garbage Detection: A Robust Deep Learning Approach

R. Regin[1,*], M. Arjun Raj[2], M. A. Thinesh[3], S. S. Mukhil Varmann[4], M. Mohamed Thariq[5], Mykhailo Paslavskyi[6]

[1,2,3,4]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[5]Department of Computer Science and Engineering, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
[6]Department of Computer Science, National Forestry University of Ukraine, Lviv, Lviv Oblast, Ukraine.
regin12006@yahoo.co.in[1], am0306@srmist.edu.in[2], tm9045@srmist.edu.in[3], sm7225@srmist.edu.in[4], thariq10410@gmail.com[5], mykhailo.paslavskyi@nltu.edu.ua[6]

Abstract: Waste generation attributable to urban life, changed consumer behaviour, and increased use of disposable lifestyles, among other contributing factors, have emerged as a major environmental and public health challenge. Improper waste management leads to pollution of land, air, and water, thereby endangering both ecosystems and communities. Slow, error-prone, and less efficient, Lesly's slower speed is incapable of achieving efficient results in traditional manual sorting methods. To address this, we propose a novel AI-based solution in this paper: a YOLOv8-based garbage detection model enhanced with a GhostNet head to further improve speed and accuracy. YOLOv8 is well-known for its real-time object detection capabilities, while GhostNet offers a lightweight yet powerful architecture that reduces processing load without compromising features. Trained on a wide variety of waste-detection datasets and complemented with rich data augmentation, this system is designed for real-world environmental conditions. It achieves impressive precision (91%), recall (88%), F1-Score (89%), and mAP@50 (0.936), proving it is well-suited for adoption in smart waste management systems. This would not only provide higher accuracy in sorting recyclables but also result in less landfilling, resource conservation, and more sustainable practices. Our research sheds light on how deep learning can be transformative in addressing waste-related issues and points toward the future, when technology will help create a cleaner, healthier planet.

Keywords: Waste Management; Garbage Detection; Real-Time Object; Urban Life; Deep Learning; Object Detection; AI in Waste Management; Data Augmentation.

Received on: 01/09/2024, **Revised on:** 15/11/2024, **Accepted on:** 20/12/2024, **Published on:** 05/06/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSCS

DOI: https://doi.org/10.69888/FTSCS.2025.000432

Cite as: R. Regin, M. A. Raj, M. A. Thinesh, S. S. M. Varmann, M. M. Thariq, and M. Paslavskyi, “AI-Powered Garbage Detection: A Robust Deep Learning Approach,” *FMDB Transactions on Sustainable Computing Systems*, vol. 3, no. 2, pp. 68–83, 2025.

1. Introduction

Waste generation is a result of human civilisation, and handling it has emerged as one of the most important environmental issues. Rapid growth in society, population, and the number of consumer goods has led to an increase in waste output. This

*Corresponding author.

increase in waste output is primarily due to changes in product consumption habits brought about by urban life and higher incomes, which have led to a greater use of disposable products and conveniences. The careless use of single-use plastics and unnecessary product packaging are among the factors driving an increase in waste output. Wastes that are not disposed of properly have devastating effects on the Earth's ecosystem. Landfills, the most common method of waste disposal, occupy a large amount of land and contribute to environmental degradation. Furthermore, badly managed garbage more often enters the water bodies, contaminating both freshwater and marine ecosystems. Plastic garbage in particular has received widespread concern for its damaging effects on marine life. Poor management of garbage not only harms the environment but also poses health risks. Open garbage dumps and open burning of garbage are common practices in underdeveloped countries with poor garbage collection and disposal facilities. These methods release harmful substances into the earth's atmosphere, land, and water, posing health risks to residents and damaging the ecosystem. People who live near such locations commonly suffer from respiratory problems, skin diseases, and other health issues. Additionally, amateur waste pickers who search for recyclable goods often work in hazardous conditions. These problems underscore the urgent need for effective waste management solutions to mitigate environmental damage, protect public health, and create a cleaner, safer future for generations to come. Proper garbage disposal is a crucial component of sustainable waste management, ensuring that waste is handled and disposed of in a manner that minimises its environmental impact. Garbage disposal keeps unwanted waste out of landfills by sorting garbage effectively, thereby minimising the amount dumped there. Efficient disposal is also crucial for conserving energy and raw materials.

Furthermore, sorting organic waste products helps us use them to produce nutrient-rich compost, which acts as a natural fertiliser. Garbage detection plays a crucial role in modern waste management, helping to address issues caused by the generation of waste. Using accurate garbage detection makes it easier to separate recyclable materials from non-recyclable ones, helping to reduce contamination in recycling streams and ensuring that resources are restored efficiently, with only waste sent to landfills. This not only conserves natural resources but also saves energy and money in the production of new materials and reduces the workload on people for sorting these wastes. Humans are prone to errors when sorting waste due to multiple factors; using garbage detection can help reduce these errors. Garbage detection also reduces the reliance on landfills. Manual sorting of garbage is time-consuming and prone to error, leading to the adoption of automated technologies. Garbage disposal is becoming more efficient thanks to AI-based models. Deep learning, a subset of AI, is used to tackle complex problems across various fields and to enhance garbage detection efficiency. Its ability to learn from large volumes of data, process image data, and recognise complex patterns makes it ideal for garbage detection. Deep learning methods utilise neural networks that can process visual data with high accuracy, enabling the efficient identification and sorting of waste. Garbage disposal is becoming more efficient thanks to AI-based models. Deep learning, a subset of AI, is used to tackle complex problems across various fields and to enhance garbage detection efficiency. Its ability to learn from large volumes of data, process image data, and recognise complex patterns makes it ideal for garbage detection. Deep learning methods utilise neural networks that can process visual data with high accuracy, enabling the efficient identification and sorting of waste. One of the primary benefits of using deep learning methods for garbage detection is its versatility.

Deep Learning models can constantly improve their accuracy by training on new data. This adaptability is crucial in waste management, where different types of waste can take various forms and require distinct handling methods. Deep learning models can learn to recognise not only ordinary images but also complex images because they utilise a neural network, enabling them to process intricate visual data. The YOLO (You Only Look Once) Model is a widely used deep learning model for object detection. The YOLO model provides a good balance between accuracy and speed when processing image data, making it a popular choice for real-time applications. The YOLOv8 incorporates features such as enhanced data augmentations and a more efficient backbone, which improve the model's precision. These features make YOLOv8 a great choice for object detection in image data. In this paper, we enhance the capabilities of the YOLOv8 model by incorporating a GhostNet-based head into the model architecture. GhostNet, with its lightweight yet powerful design, achieves high efficiency and reduces image processing time while maintaining high accuracy. Using a GhostNet-based head, our customised YOLOv8 model excels at detecting various types of garbage with high accuracy. We have utilised the Waste Detection Dataset, which encompasses a variety of waste types. Data augmentations were used during training to improve the model's capacity to generalise across various situations. The proposed YOLOv8 GhostNet Model has achieved a precision of 91%, a recall of 88%, an F1-score of 89%, and mAP@50 of 0.936. This high precision and recall show that the model is well-trained and can be used in real-time applications [16]. This paper highlights the significance of technology in addressing waste generation and underscores the need for ongoing research and development in this field. As urbanisation and consumption habits grow, more sustainable waste management methods should be implemented to create a healthier planet for future generations.

1.1. Objective

- To develop an AI-powered Garbage detector that uses a deep learning algorithm to detect and classify various forms of garbage.
- To optimise model training, a large collection of garbage photos will be preprocessed.

- To achieve real-time or near-real-time detection, enabling more efficient waste management.
- To enhance the model's efficiency, a GhostNet-based head is incorporated into the YOLOv8 model, resulting in increased processing speed without compromising accuracy.

2. Review of Literature

Wu et al. [1] address the challenges in classifying household waste images, which often feature complex backgrounds, variable lighting, diverse angles, and varying shapes. To advance research in this domain, the authors introduce the 30 Classes of Household Garbage Images (HGI-30) dataset, comprising 18,000 images across 30 distinct categories of household garbage. This publicly available dataset is designed to facilitate the development of accurate and robust methods for garbage recognition. Additionally, the paper presents experimental analyses of state-of-the-art deep convolutional neural network (CNN) methods on HGI-30, providing baseline results for future studies. Ma et al. [2] present the L-SSD algorithm, an enhanced Single Shot Multibox Detector (SSD) tailored for intelligent trash classification and recognition. Addressing challenges such as the small size of waste items and low-resolution images, L-SSD incorporates a novel feature fusion module that combines features from multiple layers and scales, forming a new feature pyramid via downsampling blocks. This design significantly improves garbage detection performance. To address the imbalance between positive and negative samples, the authors employ Focal Loss with balanced cross-entropy, emphasising training on more challenging and meaningful samples. Additionally, they replace the traditional VGG16 backbone with ResNet-101 to enhance detection accuracy and adopt Soft-NMS (Non-Maximum Suppression) to better handle overlapping detection boxes by suppressing, rather than eliminating, less desirable boxes. Experimental results demonstrate that L-SSD outperforms many state-of-the-art object detection algorithms in both accuracy and speed.

Zeng et al. [3] proposed an innovative approach to environmental monitoring by utilising airborne hyperspectral imaging (HSI) for large-area waste detection. Recognising the lack of publicly available hyperspectral garbage datasets, the authors developed and released the Shandong Suburb Garbage dataset, providing a valuable resource for future research in this domain. To effectively classify pixels within HSI data, the study proposes the Multi-Scale Convolutional Neural Network (MSCNN). This network is designed to generate binary garbage segmentation maps, enhancing the accuracy of garbage detection across expansive regions. Following segmentation, the authors employ Selective Search, an unsupervised region proposal generation algorithm, along with Non-Maximum Suppression (NMS) to determine the precise locations and sizes of garbage areas identified in the segmentation maps. Huang et al. [4] provide an innovative approach to detecting underwater debris, which is crucial for preserving marine environments. Leveraging the YOLO (You Only Look Once) object detection framework, the authors have developed a model tailored for the complexities of underwater imagery, characterised by low visibility and diverse debris types. The YOLO-MES model emphasises a lightweight architecture, ensuring efficient deployment on resource-constrained devices typically used in marine monitoring. Experimental results demonstrate that YOLO-MES achieves high detection accuracy while maintaining real-time processing capabilities, making it a practical solution for large-scale underwater garbage detection and contributing significantly to marine conservation efforts.

Rahman et al. [5] present an IoT-enabled intelligent garbage management system designed for smart cities, emphasising fairness in waste collection. Using LoRa technology for real-time monitoring, the system optimises garbage collection routes and schedules while ensuring equitable service distribution. The study highlights cost efficiency, scalability, and improved urban waste management through data-driven decision-making. The proposed framework enhances sustainability by reducing operational costs and environmental impact. Alsuwaylimi [6] discusses YOLOv8-Seg for real-time underwater debris detection using instance segmentation. It improves object recognition in challenging underwater environments, leveraging the TrashCan dataset for training. The model incorporates advanced feature extraction techniques to enhance accuracy and efficiency, making it well-suited for real-time monitoring of marine pollution. Experimental results show superior performance in detecting submerged waste compared to existing methods. Bai et al. [7] explore lightweighting the YOLOv7 model using group-level pruning to enhance computational efficiency while maintaining detection accuracy. It introduces a structured pruning method that leverages network dependence relationships and channel pruning to optimise model size. The proposed approach integrates an attention mechanism to refine feature extraction, making the model more suitable for real-time object detection, including garbage classification. Experimental results demonstrate significant reductions in model complexity with minimal loss of performance. He et al. [8] explore EC-YOLOX, a deep-learning algorithm designed for detecting floating objects in complex water environments. It likely builds upon YOLOX with enhancements such as improved feature extraction, an attention mechanism, and a refined loss function to reduce missed detections.

The method aims to handle challenging scenarios, such as reflections and occlusions, in river and lake environments. He et al. [9] present an improved YOLOv3 model for monitoring waste collection and transportation. The enhancements are likely to include triplet attention for improved feature extraction and depth-wise separable convolution for computational efficiency. The model aims to improve real-time tracking and classification of waste transport vehicles in urban environments. Abdu and Mohd Noor [10] provide a comprehensive analysis of deep learning applications in waste detection and classification. It critically evaluates various image classification and object detection models, highlighting their roles in addressing waste management challenges. A significant contribution of this work is the compilation of over twenty benchmarked trash datasets, offering a valuable resource for researchers in the field. The study also discusses the limitations of current methodologies. It outlines potential directions for future research, aiming to enhance the efficiency and effectiveness of waste management systems by applying advanced deep learning techniques.

Sheng et al. [11] review an advanced IoT-based smart waste management system that leverages LoRa communication and TensorFlow's deep learning capabilities to revolutionise waste segregation and monitoring. This system features a multi-compartment bin that autonomously sorts waste into categories such as metal, plastic, paper, and general waste, using servo motors for precise compartmentalisation. A camera integrated with a Raspberry Pi 3 Model B+ captures images of disposed items, which are then processed by a TensorFlow-trained object detection model to accurately classify the waste. Ultrasonic sensors monitor the fill levels of each compartment, while a GPS module tracks the bin's location and timestamp. Data regarding bin status, including location and fill levels, is transmitted via the energy-efficient LoRa protocol, ensuring long-range communication with minimal power consumption. Additionally, an RFID module is incorporated to authenticate waste management personnel. This holistic approach not only enhances operational efficiency but also promotes effective source segregation of waste, contributing significantly to sustainable urban waste management practices. Sallang et al. [12] explore an innovative IoT-driven smart waste management system that integrates TensorFlow Lite and the LoRa-GPS Shield to enhance waste segregation and monitoring. Utilising the SSD MobileNetV2 Quantised model, the system is trained to classify waste such as paper, cardboard, glass, metal, and plastic. A camera module connected to a Raspberry Pi 4 captures images of disposed items, which are then processed by the TensorFlow Lite model for real-time waste classification. Based on the classification, a servo motor directs the waste into the appropriate compartment within the bin.

Ultrasonic sensors monitor the fill levels of each compartment, while a GPS module records the bin's real-time location. Data on bin status, including location and fill levels, is transmitted via the LoRa communication protocol, enabling efficient waste collection planning and management. Additionally, an RFID-based locking mechanism ensures that only authorised personnel can access the bin for maintenance, enhancing security and operational efficiency. Zhou et al. [13] provide SWD, an advanced object detection model tailored for identifying solid waste in aerial imagery. The model incorporates an Asymmetric Deep Aggregation (ADA) network, which utilises structurally reparameterized asymmetric blocks to effectively extract features from waste materials that often exhibit subtle visual characteristics. To address challenges posed by blurred boundaries in aerial images, the study presents the Efficient Attention Fusion Pyramid Network (EAFPN), which enhances contextual and multi-scale geospatial information via attention fusion. The researchers compiled the Solid Waste Aerial Detection (SWAD) dataset, consisting of aerial images from Henan Province, China, to evaluate the performance of SWDet. Experimental results demonstrate that SWDet surpasses existing methods in detecting solid waste within aerial images. The authors have made the code publicly accessible on GitHub for further research and development.

Wang et al. [14] provide a comprehensive analysis of deep learning applications in waste monitoring, utilising unmanned aerial vehicles (UAVs) and satellite imagery. It systematically examines remote sensing datasets pertinent to solid waste and marine debris detection, detailing nine publicly available datasets and their respective applications. The paper categorises monitoring methods into semantic segmentation and object detection, offering insights into pixel-level classification and object-level localisation. Benchmark results from recent studies are summarised to evaluate the performance of various approaches. The authors discuss current limitations in the field and propose future research directions to enhance the efficacy of waste monitoring systems. Silva et al. [15] present a cutting-edge method for the automated detection of Chilean mine waste storage facilities (MWSFs) using Sentinel-2 satellite imagery and advanced deep learning models. The research focuses on waste rock dumps (WRDs) and leaching waste dumps (LWDs), which are prevalent in Chile's mining sector. A significant contribution of this work is the development of MineWasteCL_DB, a comprehensive public dataset comprising over 30,000 annotated images and 320,093 labels for various MWSF types, including tailings storage facilities (TSFs), WRDs, and LWDs. The study employs the YOLOv8x-seg model, selected for its high precision, to validate the presence of 96.15% of officially registered TSFs. This approach enhances the efficiency of MWSF detection and provides a scalable solution for monitoring and managing mining waste, thereby supporting environmental management and regulatory compliance in the mining industry.

3. Proposed Methodology

This research presents a more effective method for identifying garbage by modifying the YOLOv8 model and replacing its standard head with a GhostNet-based head. We aim to maintain highly accurate detection while keeping computations smooth and efficient. Our method follows a clear plan that covers the following key steps: gathering data, preparing it, refining the model, training it, and evaluating its performance. The first step is to create a comprehensive dataset that showcases various types of garbage, including plastic, paper, metal, glass, and food waste. Since original images differ in quality and dimensions, it's essential to prepare them in advance. This process includes resizing to a uniform standard and applying techniques to boost diversity. These techniques include flipping, adjusting the scale, modifying colours (HSV), utilising mosaic effects, and blending images. These help the model better handle real-life differences. The next step is to change the dataset into the YOLO format. In this format, objects have labels with boxes around them, which makes training more effective. Our approach centres on YOLOv8, a popular object detection system built for real-time use. We replace the standard head with a GhostNet-based one. We have a detection mechanism that's both lightweight and robust. This minor modification enables the entire process to proceed a bit faster without compromising the system's sharp eye for detail. When it comes to the training, we're tuning the switched-up YOLOv8 design over 100 training cycles. We run the training on a powerful GPU, which allows us to handle big datasets. To assess the model's effectiveness, we examine key performance indicators. These include precision, recall (mAP@50 and mAP@50-95), and F1-score. We also create a confusion matrix to see how it categorises different waste types and to identify any patterns where it makes errors.

3.1. YOLOv8

YOLOv8 is the cutting edge of the "You Only Look Once" (YOLO) series by Ultralytics that was incepted by this name. YOLOv8 is a completely new object detection and classification library built upon YOLO, delivering accuracy, speed, and flexibility for real-time use cases. YOLO models differ from traditional object detection pipelines, which have multiple steps, by performing the entire process in a single step. This minimal process beats both in terms of computational efficiency and provides a better and faster response. Deeper yet, YOLOv8 builds on this, updating the existing framework with modern conveniences needed for current AI jobs. YOLOv8 features notable architectural improvements over YOLO's well-established architecture. It uses a compact deep backbone, CSPDarkNet, and also PANet for feature fusion across different layers. Together, they enable faster and more accurate object detection. It also introduces an anchor-free detection runtime in YOLOv8, making it flexible for all kinds of object shapes and sizes, rather than relying on anchor boxes. One of the key features of YOLOv8 is its dynamic compute allocations. With support for this feature, the model's performance will be dynamically optimised for the targeted hardware type (edge devices with limited resources, high-performance GPUs, or cloud-based servers). YOLOv8 is super generic and can be deployed as a tool across multiple platforms and environments. YOLOv8 isn't just about spotting things; it identifies specific objects and estimates their positions, making it a versatile tool for tasks that rely on visual recognition. To teach this model, they employ techniques to help it learn faster and more effectively, such as AdamW, cosine learning rate scheduling, and slightly fuzzy labels. And when it's figuring out how well it's doing, it looks at how well it's guessing shapes, whether it's got its categories right, and if it's pretty confident it's seeing something that's there.

3.2. YOLOv8-GhostNet Head

The integration of a GhostNet-based detection head into the YOLOv8 architecture yields significant improvements in speed, precision, and computational efficiency, which can be crucial when resources are limited and real-time prediction is required. The YOLOv8 detection head is known for its outstanding precision but is computationally demanding, making it difficult to deploy on edge devices and embedded systems. The GhostNet-head offers a lighter approach by utilising GhostBottleneck layers. These layers are designed to generate additional feature maps with less expensive linear operations rather than the usual intensive convolutional computations. YOLOv8 GhostNet-based head architecture decreases computational burden while maintaining the model's ability to detect and classify images precisely. The detection head replaces two Convolutional layers with two GhostBottleneck layers. This channel reduces the number of layers from 169 to 92, speeding up the processing pipeline while maintaining accuracy. The model performs well even on low-power hardware by leveraging GhostNet's architecture, which is well known for its efficiency; this makes it ideal for real-time garbage detection. To validate the performance of the improved architecture, we trained the YOLOv8-GhostNet model on a diverse dataset for garbage detection, which contains multiple waste categories, including batteries, biological waste, brown glass, cardboard, clothes, green glass, metal, plastic, paper, shoes, trash, and white glass. The training process involves data pre-processing and the use of augmentation techniques such as HSV colour space transformation, horizontal flipping, scaling, Mosaic, and Mixup. These methods are applied to improve the model's generalisation ability. We trained the model for over 100 epochs using the AdamW optimiser and evaluated using precision, recall, and F1-score.

3.3. Architecture Diagram

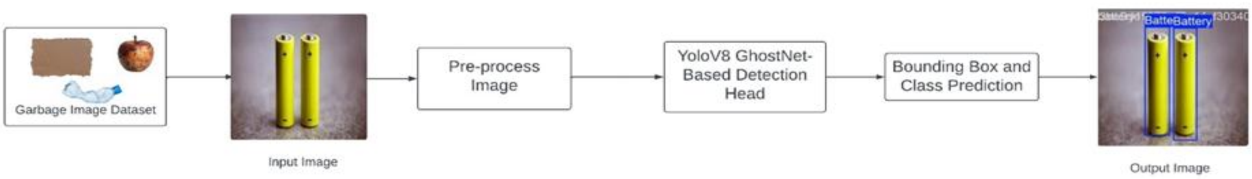

Figure 1: General architecture diagram of the garbage detection system

Figure 1 shows how it works: a garbage detection pipeline based on the Yolov8 GhostNet model and the Garbage Image Dataset (categories of garbage: organics, plastics, metal, organic waste, batteries). It preprocesses the input image (resizing, normalisation, and augmentation for model generalisation). The image is then passed to the YOLOv8 model, which uses a GhostNet backbone for both feature extraction and computational reduction. The model predicts the bounding box and its label to determine whether it is a battery, plastic, or other waste material. The output image displays the detected object with labels for class and confidence score, facilitating easy and accurate garbage categorisation.

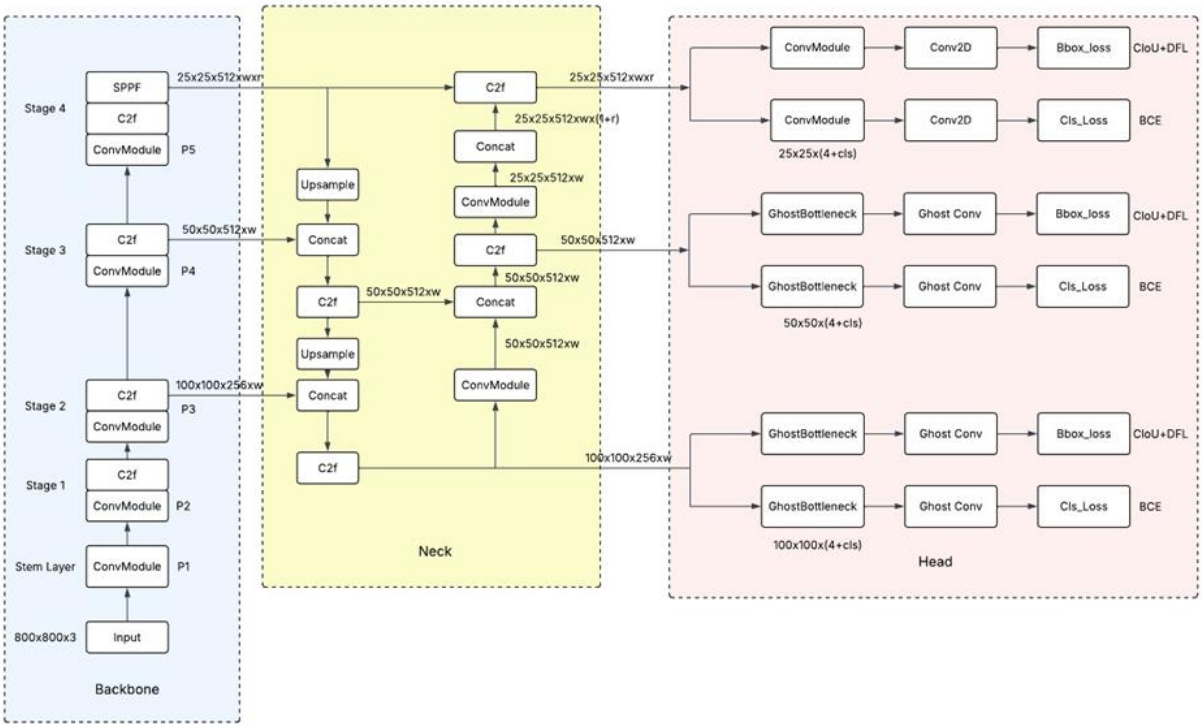

Figure 2: Architecture diagram of the Yolov8 ghostnet-based model

Figure 2 depicts the Architecture of the modified YOLOv8, including a detection head based on GhostNet to reduce computational demands while maintaining effective feature extraction. It has three main parts, Backbone: convolution-layers and C2f blocks at multiple stages, P2 to P5 (output dimensions are 100×100×256 –> 25×25×512) which extract hierarchical features; Neck: (upsampling, concatenation and feature refinement with C2f and CNNs) doing feature combination across scales useful to multi-scale detection together; Head — the standard detection layers in YOLO are attempted to be replaced by GhostBottleneck and Ghost Convolution layers to drastically decrease computation cost, while keeping accuracy. Detection head for regression and classification loss, utilising CIoU+DFL boxes in the bounding-box regression layer, while classification employs BCE to ensure accurate localisation of objects and their corresponding categories. This highly optimised structure strikes the right balance between speed and accuracy; hence, it is well-suited for real-time applications such as garbage classification and detection.

3.4. Algorithm

BEGIN

Install Required Libraries

- Install Ultralytics, scikit-learn, matplotlib, seaborn, torch, and torchvision.

Import Necessary Modules

- Import torch, numpy, pandas, matplotlib.pyplot, seaborn, cv2, and ultralytics. YOLO.

Define a Custom GhostNet-Based YOLO Head

- Create a GhostBottleneck class to efficiently extract features using depthwise convolutions.
- Implement the YOLOGhostHead class, which replaces the default YOLO detection head with a GhostNet-based structure.

Load pre-trained YOLOv8 Model.

- Initialize the YOLOv8 model using a pre-trained weight file (yolov8m.pt).
- Replace the default model detection head with the GhostNet-based detection head.

Train the YOLOv8-GhostNet Model

Specify training parameters:

- Dataset Path: Load dataset annotation file (data.yaml).
- Training Configuration:
    - Epochs = 100
    - Image Size = 800 × 800
    - Batch Size = 16
    - Number of Workers = 4
- **Data Augmentation:** Apply transformations like HSV colour shifts, rotation, translation, scaling, shearing, perspective distortion, flipping, mosaic, and mixup.
- **Optimiser and Learning Rate:** Use Cosine Learning Rate Scheduler with initial learning rate (lr0 = 0.001) and final learning rate (lrf = 0.0001).
- **Early Stopping:** Set patience = 10 to stop training when no improvement is observed.
- **Training Execution:** Train the model using the modified architecture and save results in "garbage_detection/optimized_yolo_dsc".

Load Training Results

- Read the training log (results.csv) from the output directory.
- Extract key metrics from the dataset:
    - Training and Validation Loss.
    - Precision, Recall, and mAP@50.
    - F1-Score Calculation using the formula:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Plot Performance Graphs:

- **Loss Curve:** Plot training vs. validation loss over epochs.
- **Accuracy Curve:** Compare training and validation accuracy (mAP@50).
- **F1-Score Curve:** Display F1-score progression across epochs.
- **Prediction Time Analysis (if available):** Analyse model inference time per epoch.

Visualise the Confusion Matrix:

- Load and display the confusion matrix from the results folder.
- If a normalised version exists, display it for a clearer class-wise performance evaluation.

Save Model and Performance Plots:

- Store trained YOLOv8-GhostNet weights, results, and evaluation graphs for further analysis.

END

3.5. Formulas

3.5.1. Object Detection Loss Function (YOLOv8)

The total loss function LLL in YOLOv8 is a combination of three components:

$$L = L_{box} + L_{obj} + L_{cls}$$

Where:

- L_box (Bounding Box Loss): Measures the difference between predicted and ground truth bounding box coordinates using CIoU loss:

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

Where:

- IoU (Intersection over Union):

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

- $\boldsymbol{\rho^2(b, b^{gt})}$ **:** Squared distance between the centre points of predicted and ground truth boxes.
- $\boldsymbol{c^2}$**:** Diagonal length of the smallest enclosing box covering both the predicted and ground truth boxes.
- **v:** Aspect ratio consistency term.
- $\mathbf{L_{obj}(ObjectnessLoss)}$**:** Measures confidence of object presence using binary cross-entropy loss (BCE):

$$L_{obj} = -\sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

where y_i is the ground truth and $\hat{y}_i$ is the predicted object confidence score.

L_{cls} (Classification Loss): Measures the correctness of predicted class labels using cross-entropy loss:

$$L_{cls} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

3.5.2. Evaluation Metrics

To measure the performance of the model, the following metrics are used:

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

- Precision measures the fraction of predicted garbage objects that are actually correct.

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

- Recall measures the proportion of actual garbage objects that are correctly detected.

$$\text{F1-score} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

- F1-score balances precision and recall for better evaluation of detection performance.

$$\text{Mean Average Precision (mAP)} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{AP_i}$$

- where AP_i is the Average Precision for each class, computed as:

$$\mathrm{AP} = \int_0^1 \mathrm{P(R)\ dR}$$

Which is the area under the Precision-Recall curve?

3.6. Existing Model

The Faster R-CNN, equipped with a ResNet-50 backbone and FPN, employs a bottom-up approach at a foundational level in this study. Faster R-CNN is another two-stage object detection framework; it clusters regions that may contain objects and subsequently classifies them. It excels at object detection by striking a balance between accuracy and robustness, particularly in scenarios involving small objects and noisy, complex backgrounds (which is ideal for handling waste). The dataset used is waste classification dataset that consisted of many categories of garbage for training and evaluate this model, the classes are Battery, Biological, Brown glass, Cardboard, Clothes, Green glass, Metal, Paper, Plastic, Shoes, Trash, White glass that Faster R-CNN could create well-formed region proposals and capture the plight object features makes it compatible with different object classes, even if some classes of overlapping products or same-looking waste material. After evaluation, the model achieved a precision of 90.77%, a recall of 90.49%, and an F1-score of 90.49%. Though Faster R-CNN can be slow because it isn't real-time (due to its computational burden), its accuracy and track record make it widely used in research and production, where speed isn't more important than precision. This study aims to measure the incremental efficiency and enhancements of the YOLOv8-GhostNet variant, serving as the primary improved model, against the baseline model for simplicity in this work.

3.7. Execution

The deployment phase of the paper is carried out through the systematic application of the methodology discussed, from data preparation to model deployment. It begins with dataset preprocessing, where images are resized, normalised, and subjected to augmentation techniques such as flipping, scaling, HSV, mosaic, and mixup to enhance the model's generalizability. This is followed by formatting the dataset in YOLO annotation format, labelling object bounding boxes with care. Following this, the YOLOv8m model is set as the base detection backbone, and the default YOLO head is replaced with a GhostNet-based head that includes GhostBottleneck modules to achieve maximum computational efficiency. In training, the pre-trained model is fine-tuned for more than 100 epochs with a batch size of 16 with optimisers such as AdamW. Robustness is enhanced through data augmentation, and performance is evaluated using critical metrics such as precision, recall, mAP@50, and F1-score. Class-wise accuracy and misclassifications are analysed through a confusion matrix. The model is then tested on test data after training, and loss curves and accuracy plots are generated to track performance trends. The trained model is then used for inference, which displays the waste objects with bounding boxes and class labels to achieve accurate detection in the real world. Computational efficiency is also tested, with notable improvements in inference speed and reduced memory usage, thanks to the integration of GhostNet. Generally speaking, this approach ensures that the proposed model strikes a balance between high detection precision and low computational cost, thereby making it applicable in practical smart waste management systems.

4. Implementation

4.1. Data and Pre-Processing

Training and validating the YOLOv8 garbage detection model is also highly dependent on the dataset we are working with, as it is a key component. It sources images of the waste (plastic, paper, metal, glass, and organic waste). Images are diversified for environmental conditions. The first preprocessing step is to clean the data. I remove any low-resolution, blurry, or duplicate images, which would only increase the dataset's size. The next one is an annotation in YOLO format; the object to be detected is given with bounding boxes and class indices, enabling actual object detection. We use some preprocessing steps to improve the model's generalisation and detection, as Key. That adjusts the input shape, regardless of the data, to maintain a uniform size and normalise pixel values for stable training. Some examples of data augmentation techniques include horizontal and Vertical Flipping. Image resizing is performed on these to better render the diversity of images. Scale augmentation, HSV transformations, mosaic augmentation, and mixup are strategies for augmenting the dataset with artificially generated data points to address differences in object locations, lighting, and background clutter. They help prevent overfitting and allow the model to learn more general characteristics. Furthermore, it is split evenly into training, validation, and test datasets, allowing the model to be fairly tested across slightly skewed data distributions. The model trained on this type of orthodox preprocessing pipeline would be able to detect with the best possible accuracy, even in a real-world classification setting that is more cluttered.

4.2. Data Visualisation

Figure 3 shows the distribution of images per class in the dataset, indicating the total number of images available for each waste object. Clean and Clothes have the most, "Plastic" is a close second, and "Cardboard" has negligible samples.

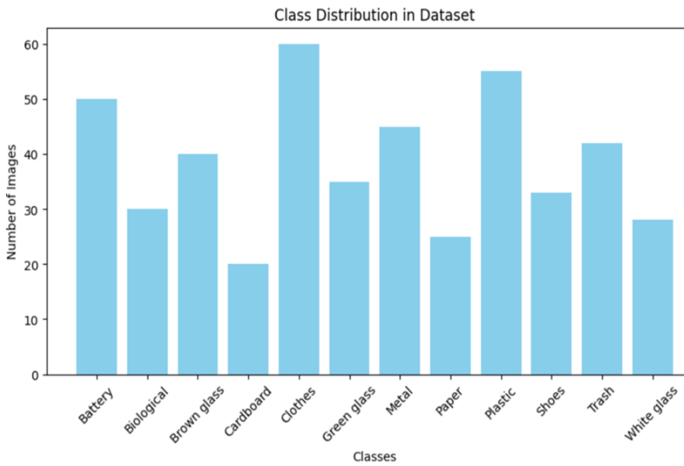

Figure 3: Class distribution in the dataset

The resulting imbalance may adversely affect model predictions, as classes may be favoured over underrepresented categories. An important requirement for learning reasonable and fair results is having a suitably balanced dataset. E.g., data augmentation for minority classes, high-level solutions not limited to these, or tweaking the training strategy with weighted loss functions for better generalisation across all categories.

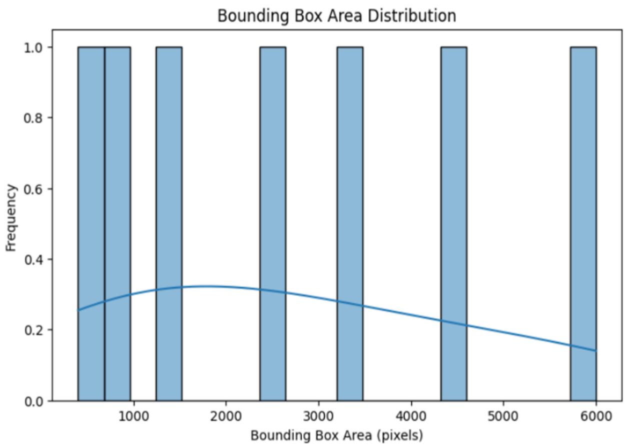

Figure 4: Bounding box area distribution

Figure 4 illustrates the distribution of bounding box areas in pixels, showing the range of object sizes in the dataset. Each bar corresponds to the number of bounding boxes of different sizes, and the density curve provides a less choppy estimate of the overall distribution. The presence of different-sized holes and a random distribution of frequency points indicates an imbalanced dataset, with some classes having significantly more or fewer instances than others. This imbalance can affect the performance of the object detection model's output: smaller boxes are usually of poorer quality, and larger ones warp bounding boxes to make it easier to catch only big objects. Methods such as data augmentation, resizing, or even anchor box optimisation can be used to achieve a more uniform object size distribution and increase the model's robustness, as well as improve its generalisation performance.

Figure 5: Sample images of different classes

Figure 5 provides a compartmental overview of several waste classes, with each class represented as a distinct object. This dataset comprises the following categories: biological waste, batteries and their packaging (brown glass), clothing, plastic, white glass, green glass, paper, cardboard, shoes, and miscellaneous trash. This representation format serves as decision support to distinguish between recyclable, biodegradable/compostable, and non-recyclable materials. Each object in this image reflects reality through different materials, shapes, and textures, resulting in a variety of datasets. e.g., the brown/green glasses represent colour-based differences among different glass types, and plastic/metal showcase different materials. Biological waste, clothes, and shoes are material categories that do not belong to either the recycling group or the organic waste group. The dataset is expanded with paper, cardboard, and other trash to enrich the training of waste recognition models. This visualisation is especially good for teaching a few-shot dataset of AI-based trash-recognition systems, and one can be assured that the examples are balanced, as they often encounter. A grid view of different waste streams in one place should help human comprehension and machine learning model development for waste sorting, as well as improve manageability CVs.

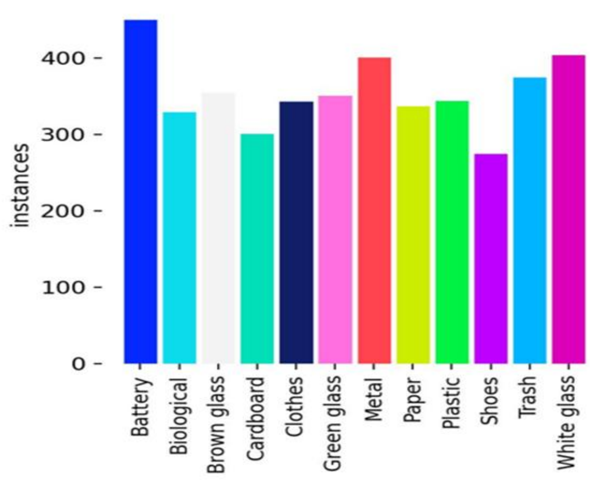

Figure 6: Instances of different classes

Figure 6: The frequency of Different Waste Categories analysed in the dataset is shown in the bar chart, which displays the type and class distribution. The battery class is the most represented, followed by metal and white glass (a lot of stuff), indicating a strong class imbalance in this set. In contrast to the lower counts for categories such as biological waste, plastic and shoes there is higher representation to keep a wide range of classes but not so many that fitting a NN model becomes hard. Using unique colours for each class made a few things look a lot clearer and, in the end, it was easier to compare how often different waste types were being picked up. The distinct categories of brown, green, and white glass imply that glass waste has been

classified in great detail. A moderate presence of cardboard, clothes, and paper, as well as garbage, ensures the dataset has a good mix of recyclables and non-recyclables. This is crucial for AI-powered waste detection models to prevent data from being overrepresented in a class. A well-balanced dataset will enable the model to better identify different waste types, resulting in effective waste segregation and recycling.

4.3. Training

GhostNet Head on YOLOv8-Based Garbage Detection Model Training Process. This constitutes the training phase of our YOLOv8 garbage detection model with a GhostNet-based head, ensuring the device can detect/objectify correctly and quickly. This involves tuning the model to achieve precise object detection and classification of waste objects whilst maintaining computational speed. The training pipeline starts by loading the preprocessed dataset into training, validation, and test splits. A rich and adequate dataset helps the model better classify waste objects, regardless of their type or environment. The YOLOv8m architecture is initialised and used to train the model with our head, gradually replacing the default detection head with GhostNet to improve computational efficiency on the end device while maintaining state-of-the-art detection accuracy. After 100 epochs with a batch size of 16, training is performed using data augmentation methods, including flipping, scaling, and colour transforms, as well as mixup, to enhance robustness. We use AdamW as the optimiser (primarily due to its higher performance tuning), and the learning rate follows a cosine annealing schedule to facilitate smooth convergence. Key evaluation metrics for model performance, such as precision, recall, mean Average Precision, and F1 Score, are monitored in real-time. Backpropagation is used to minimise a loss function that combines classification, object-less, and localisation losses. Experiments are run on a high-performance GPU to accelerate computation and optimise performance. Furthermore, 10 epochs patience with early stopping is applied to avoid overfitting when the validation loss stops decreasing. The model's performance is then examined through validation results (lower error), loss curves, and accuracy graphs after training to ensure our model is accurate, runs faster, and is highly efficient.

5. Results and Discussion

Evaluation of the proposed YOLOv8 garbage detection model with a GhostNet-based head includes precision, recall, mean Average Precision (mAP@50, mAP@50-95), and F1-score metrics. The model has been trained on a large dataset of 12 waste categories and tested on new images outside this dataset to ensure it achieves real-world performance. As a result, the modified architecture performs accurately and efficiently for real-time waste classification, garbage management, and related applications. Based on the evaluation metrics, the model performed well, with an mAP@50 of 0.936 across our waste classes, indicating strong detection performance across categories. The evaluation metrics Table shows that our model achieved an overall mAP@50 of 0.936. This paper has achieved high Precision for each garbage type. The highest-accuracy categories are Brown Glass (mAP@m00 = 0.098) and clothes. The confusion matrix indicates that mislabelling primarily occurs among visually similar waste items — specifically, glass and plastic materials — suggesting that collecting more images and refining this dataset could lead to more accurate classifications.

The GhostNet-based head made a significant contribution to inference speed and model size reduction, albeit at the expense of some detection accuracy, as assessed by IoU. GhostBottleneck modules help the model to be more computationally lightweight, making it suitable for deployment on edge devices and low-power systems. As both curves (training and validation) in this graph are stable, it indicates good convergence, suggesting that very little overfitting has occurred, given the extensive data augmentation and adaptive learning rate schedule. This study shows that the proposed approach for improving waste detection and classification efficiency is effective. However, future work could potentially improve this by adding more lightweight architectures to the list, hyperparameter tuning, or using more sophisticated post-processing to fine-tune detections. This in-depth discussion of the findings offers critical insights into the trade-offs among model accuracy, throughput, and ease of learning and direct deployment, thereby making new smart waste solution proposals more efficient and scalable.

Table 1: Comparison of faster R-CNN and yolov8-ghostnet

Metric	Faster R-CNN (ResNet-50)	YOLOv8-GhostNet
Precision	90.77%	91%
Recall	90.49%	88%
F1-Score	90.49%	89%

Table 1 compares Faster R-CNN and YOLOv8-GhostNet. Faster R-CNN achieves precision (90.77%) and recall (90.49%) thanks to its region proposal network (RPN), which produces well-defined bounding boxes, including those for smaller or overlapping objects. The downside is that it operates in two stages (first a region proposal, then classification), which results in a significant compute time lag. That time lag makes it unsuitable for any applications requiring real-time response, which is especially important for automated waste sorting.

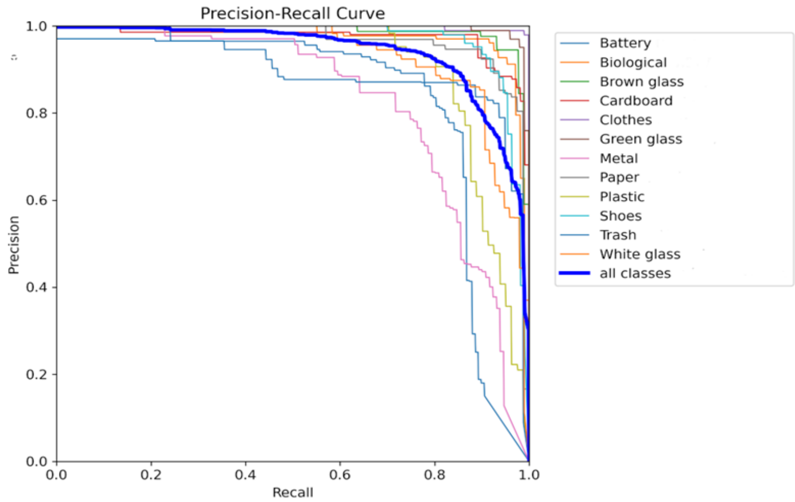

Figure 7: Precision-recall curve

Figure 7 and Table 2 evaluate the YOLOv8-GhostNet model's performance in detecting various types of garbage, as illustrated in the Precision-Recall (PR) curve and Mean Average Precision (mAP) values. The PR curve shows the precision-recall trade-off for all waste categories, with a higher curve characterising more accurate classification. The solid blue curve shows the global model's performance across all classes, suggesting a well-balanced dice (precision and recall).

Table 2: Tabulated values of the precision-recall curve

Waste Category	mAP@0.5
Battery	0.830
Biological	0.977
Brown Glass	0.983
Cardboard	0.971
Clothes	0.993
Green Glass	0.991
Metal	0.824
Paper	0.963
Plastic	0.909
Shoes	0.969
Trash	0.897
White Glass	0.928
All Classes (mAP@0.5)	0.936

Categories: Clothes, Green Glass, and Brown Glass have nearly 100% detection accuracy; Metal and Trash have lower metrics, indicating some mislabeling. On the right is the Table of mAP@0.5 values for each class, with the overall model reaching a high detection accuracy of 0.936.

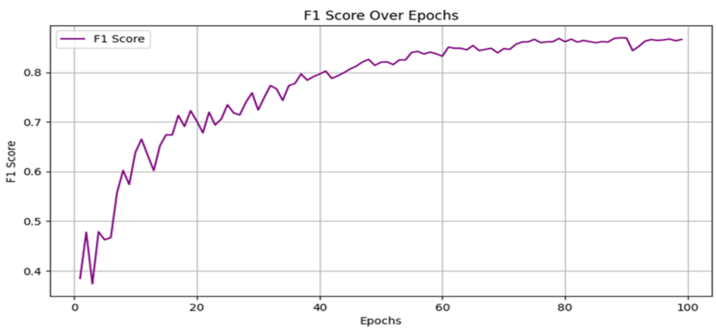

Figure 8: F1 score over epochs

The highest results are achieved in clothes (0.993), Green Glass (0.991), and Brown Glass (0.983), with Metal (0.824) and Trash (0.897) performing the worst. This means the model still has room to improve in classifying more challenging categories, but it performs quite well for most types of waste. The YOLOv8-GhostNet model demonstrates significant potential for smart waste management applications, accurately classifying various types of waste. Figure 8 and Table 3 illustrate the performance of our modified YOLOv8 network, which was trained using a GhostNet-based head. The F1-score begins relatively low in the top-left corner, indicating that the model initially struggles to distinguish between the waste classes. However, as we sustain the training, we observe the score jumps up rapidly, indicating improvement in precision and recall around epoch 40.

Table 3: Tabulated values of F1 score over epoch

Epochs	F1-Score
0	0.35
10	0.55
20	0.72
30	0.75
40	0.78
50	0.81
60	0.83
70	0.85
80	0.86
90	0.87
100	0.88

The model then stabilises, and by consistently reaching the 60th F1 above a threshold of 0.85, the classification performance can be considered satisfactory. This means that the model is mastering different waste categories fairly well, with almost zero misclassification for all categories. The flatness of the curve after the 80th epoch suggests that further training yields marginal returns, indicating convergence in this model. The high F1-score for whole-detection, combined with the GhostNet-based modification, demonstrates that the model is both fast and efficient in terms of computational cost and accuracy.

Figure 9: F1 score over epochs

Figure 9 and Table 4 show the model's training convergence. The training loss is in blue, and the validation loss (red) is plotted over 100 epochs. They both start high but decrease over time with training, which is beneficial: effective learning. The validation loss behaves similarly to the training loss, suggesting the model is generalising well to unseen data and not overfitting. There should be a slight difference between the training and validation losses.

Table 4: Tabulated values of F1 score over epoch

Epochs	Training Loss	Validation Loss
0	2.1	1.8
10	1.2	1.0
20	1.0	0.85
30	0.9	0.75
40	0.85	0.72
50	0.8	0.7

60	0.75	0.68
70	0.72	0.66
80	0.7	0.65
90	0.65	0.62
100	0.5	0.6

However, as we can see, they are not far apart, even across epochs, indicating that the model still exhibits good generalisation. A steady decrease in loss towards the last epochs shows adaptive learning rate adjustments that were helping for optimisation, quite sharp for sure. The reduced loss values as a whole still indicate the model works and is stable for our garbage classification.

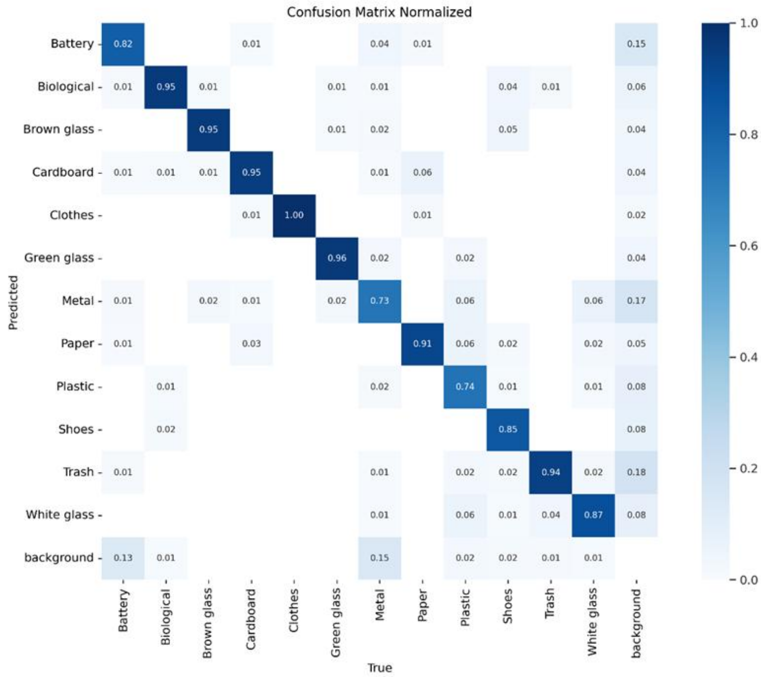

Figure 10: Confusion matrix of the yolov8-ghostnet model

Figure 10 clearly illustrates how robust the classification model performs in recognising distinct types of waste. When "Clothes", the model has perfect accuracy and excels at things like "Green glass" (96%), "Biological" (95%), "Brown glass", and even "Cardboard" (95%), which is a testament to its discriminating power. Moreover, other parts of "Data" -- such as "Paper" (91%), "Trash" (94%), and "White glass" (87%) (>80% accuracy) are indicative that the model is trustworthy. Robustness is evident even when accuracy is slightly degraded, with misclassifications remaining low—more than enough to be considered robust. The refined colour gradient makes the model easy to interpret, highlighting high-confidence regions and showing where the model really shines. This is a great visualisation that shows how effective your model is and offers a taste of real-world applications (such as waste sorting or recycling automation).

6. Conclusion

The proposed approach for garbage detection combines a YOLOv8 model with a GhostNet-based head to improve the model's efficiency. By utilising YOLOv8's object detection framework and the GhostNet-based head, the model ensures precise detection across various types of waste. By leveraging GhostNet's lightweight yet effective capabilities, the model reduces processing time while maintaining accuracy. The model is trained on a wide range of garbage image datasets to improve generalisation across different environments and waste types. The model has achieved a precision of 91%, a recall of 88%, an F1 score of 89%, and a mAP@50 of 0.936. The proposed approach shows significant improvements in garbage detection accuracy. Future work may include expanding the dataset to include more types of garbage and testing deployment in a real-world system.

Acknowledgement: The authors express their sincere gratitude to all authors who have contributed equally to the completion of this work.

Data Availability Statement: The data and materials used in this research on AI-powered garbage detection: a robust deep learning approach are available upon reasonable request to the corresponding author for academic and research purposes.

Funding Statement: This research and manuscript preparation were completed without any external funding or financial assistance.

Conflicts of Interest Statement: The authors collectively declare that there are no conflicts of interest related to this research, its results, or its publication.

Ethics and Consent Statement: All ethical standards were adhered to during the research process. Necessary permissions and informed consent were obtained from the respective organisation and participants involved in data collection.

References

1. Z. Wu, H. Li, X. Wang, Z. Wu, L. Zou, and L. Xu, "New Benchmark for Household Garbage Image Recognition," *Tsinghua Science and Technology*, vol. 27, no. 5, pp. 793–803, 2022.
2. W. Ma, X. Wang, and J. Yu, "A Lightweight Feature Fusion Single Shot Multibox Detector for Garbage Detection," *IEEE Access*, vol. 8, no. 10, pp. 188577–188586, 2020.
3. D. Zeng, S. Zhang, F. Chen, and Y. Wang, "Multi-Scale CNN Based Garbage Detection of Airborne Hyperspectral Data," *IEEE Access*, vol. 7, no. 7, pp. 104514–104527, 2019.
4. C. Huang, W. Zhang, B. Zheng, J. Li, B. Xie, and R. Nan, "YOLO-MES: An Effective Lightweight Underwater Garbage Detection Scheme for Marine Ecosystems," *IEEE Access*, vol. 13, no. 3, pp. 60440–60454, 2025.
5. M. A. Rahman, S. W. Tan, A. Taufiq Asyhari, I. F. Kurniawan, M. J. F. Alenazi, and M. Uddin, "IoT-Enabled Intelligent Garbage Management System for Smart City: A Fairness Perspective," *IEEE Access*, vol. 12, no. 6, pp. 82693–82705, 2024.
6. A. A. Alsuwaylimi, "Enhanced YOLOv8-Seg Instance Segmentation for Real-Time Submerged Debris Detection," *IEEE Access*, vol. 12, no. 8, pp. 117833–117849, 2024.
7. X. Bai, H. Shi, Z. Wang, H. Hu, and C. Zhang, "Study of YOLOv7 Model Lightweighting Based on Group-Level Pruning," *IEEE Access*, vol. 12, no. 7, pp. 96138–96149, 2024.
8. J. He, Y. Cheng, W. Wang, Y. Gu, Y. Wang, and W. Zhang, "EC-YOLOX: A Deep-Learning Algorithm for Floating Objects Detection in Ground Images of Complex Water Environments," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, no. 2, pp. 7359–7370, 2024.
9. Y. He, J. Li, S. Chen, Y. Xu, and J. Liang, "Waste Collection and Transportation Supervision Based on Improved YOLOv3 Model," *IEEE Access*, vol. 10, no. 8, pp. 81836–81845, 2022.
10. H. Abdu and M. H. Mohd Noor, "A Survey on Waste Detection and Classification Using Deep Learning," *IEEE Access*, vol. 10, no. 12, pp. 128151–128165, 2022.
11. T. J. Sheng, M. S. Islam, N. Misran, M. H. Baharuddin, H. Arshad, and Md. R. Islam, "An Internet of Things Based Smart Waste Management System Using LoRa and Tensorflow Deep Learning Model," *IEEE Access*, vol. 8, no. 8, pp. 148793–148811, 2020.
12. N. C. A. Sallang, M. T. Islam, M. S. Islam, and H. Arshad, "A CNN-Based Smart Waste Management System Using TensorFlow Lite and LoRa-GPS Shield in Internet of Things Environment," *IEEE Access*, vol. 9, no. 11, pp. 153560–153574, 2021.
13. L. Zhou, X. Rao, Y. Li, X. Zuo, Y. Liu, and Y. Lin, "SWDet: Anchor-Based Object Detector for Solid Waste Detection in Aerial Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 16, no. 11, pp. 306–320, 2023.
14. B. Wang, Y. Xing, N. Wang, and C. L. P. Chen, "Monitoring Waste from Uncrewed Aerial Vehicles and Satellite Imagery Using Deep Learning Techniques: A Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, no. 11, pp. 20064–20079, 2024.
15. M. Silva, G. Hermosilla, G. Villavicencio, G. Cocca-Guardia, P. Breul and V. Aprigliano, "Automated Detection of Chilean Mine Waste Storage Facilities Using Advanced Deep Learning Models and Sentinel-2 Satellite Imagery," *IEEE Access*, vol. 13, no. 2, pp. 39666–39679, 2025.
16. SAFITRI, "Waste Detection Dataset," *Roboflow Universe, Roboflow*, 2023. AvaIlabe: https://universe.roboflow.com /safitri/waste-detection-lvtny [Accessed by 29/07/2024].