Waste Object Detection and Classification

Group 8 - Amedeo Filippini, Bumin Kagan Cetin, Daniel Mitoiu

1 Introduction

The rapid increase in global waste production poses significant environmental and logistical challenges. Accurate waste identification and classification are crucial for efficient recycling processes and environmental cleanup. Traditional manual sorting methods are not only labor-intensive but also highly prone to human error, resulting in inefficiencies. Additionally, the heterogeneous nature of waste materials complicates manual sorting, further underscoring the need for automated systems.

In this work, we aim to develop a system for detecting and classifying waste objects in images using advanced computer vision and deep learning techniques. The system leverages YOLO (You Only Look Once) for object detection and a ResNet-18 convolutional neural network (CNN) for classification into 60 material super-categories such as plastic, metal, glass, and paper. This research introduces a scalable and efficient pipeline for addressing the inefficiencies inherent in waste management systems.

2 Problem Formulation

The task involves two primary challenges: object detection and material classification. Object detection requires the identification and localization of waste objects in images, achieved by predicting bounding boxes. Material classification assigns each detected object to a specific category based on its material composition.

This problem has profound implications for modern recycling systems. Automated waste classification can reduce contamination in recycling streams, improve material recovery rates, and lower operational costs. From an environmental standpoint, the deployment of such systems can minimize landfill usage and conserve natural resources through enhanced recycling. The YOLO object detection framework was chosen due to its ability to achieve real-time detection while maintaining high accuracy, making it suitable for deployment in dynamic, real-world environments.

3 Contributions

This study makes several contributions to the domain of automated waste detection and classification:

- We employ the TACO (Trash Annotations in Context) dataset, a publicly available collection of annotated images of waste in real-world contexts.
- A fine-tuned YOLOv8 model is implemented to detect waste objects.
- A classification pipeline using ResNet-18 is developed to categorize detected objects by material type.
- A structured pipeline for preprocessing, visualization, and evaluation of waste detection and classification tasks is proposed.

4 Background and Related Works

Object detection has seen significant advancements in recent years, with YOLO emerging as one of the most effective frameworks. Unlike traditional multi-stage methods such as R-CNN, YOLO adopts a single-stage approach, predicting bounding boxes and class probabilities simultaneously using a convolutional neural network. The YOLOv8 version used in this project incorporates improvements such as CSPNet (Cross-Stage Partial Networks) for feature extraction and a Feature Pyramid Network (FPN) for multi-scale detection, ensuring a balance between computational efficiency and accuracy.

5 Methodology

5.1 Data Collection and Preprocessing

The TACO dataset was accessed via its official GitHub repository, and the annotations were parsed from JSON files. Images were downloaded, and corresponding bounding box information was extracted. These images were organized into structured directories for training, validation, and testing.

To ensure consistency and reliability, annotations were visually validated using Python libraries such as Matplotlib. During this process, we observed that some images contained multiple objects with separate annotations. To address this, the object with the largest bounding box area was selected as the primary object for classification. Images with missing or incorrect annotations were excluded to enhance the overall quality of the dataset.

The dataset was then split into three subsets: training (70%), validation (15%), and testing (15%). To maintain a balanced representation across subsets, random shuffling was applied. Manual adjustments were performed to balance the dataset and ensure equitable representation of all classes.

5.2 Object Detection

The YOLOv8 model was fine-tuned on the training subset of the TACO dataset to perform waste object detection. YOLOv8 operates as a single-stage object detection framework that divides the input image into a grid of cells. For each cell, it simultaneously predicts bounding box coordinates, objectness scores, and class probabilities.

The performance of YOLOv8 was evaluated using the Intersection over Union (IoU) metric, defined as:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

The model achieved a mean IoU of 68%, with 74% of the predictions exceeding the 0.5 IoU threshold.

5.3 Classification

For the classification task, a ResNet-18 model was developed to process cropped image regions corresponding to the detected bounding boxes generated by the YOLOv8 model. These cropped regions represent individual waste objects, which are further classified into predefined material categories (e.g., plastic, metal, glass, paper).

The final fully connected layer of the pretrained ResNet was replaced to match the number of classes in the TACO dataset. The CrossEntropyLoss function was used as the primary loss metric, and the Adam optimizer was employed to train the model effectively.

5.4 Evaluation Metrics

The evaluation pipeline uses accuracy, precision, and recall as primary metrics for assessing performance on validation and test datasets.

6 Results

The YOLOv8 model achieved an average IoU of 0.68 on the validation set. However, the classification model achieved a low accuracy of 15% on the test set. The confusion matrix revealed very low precision across most material categories. Visualizations of detected bounding boxes and material labels confirmed that misannotations were a significant contributor to low metrics.

7 Limitations of the Project

- **Dataset Imbalance:** Rare classes remain underrepresented despite manual balancing efforts.
- **Single Object Focus:** Only the object with the largest bounding box is considered, ignoring contextual relationships.
- **Dependence on Bounding Box Quality:** Errors in bounding box annotations compromise training data quality.
- **Computational Constraints:** The use of computationally intensive architectures limits deployment in resource-constrained environments.
- **Focus on Classification:** Bounding box annotations are only used for preprocessing, limiting detection and localization capabilities.

8 Discussion

The results demonstrate the potential of the proposed system for real-world waste management applications. Expanding the dataset to include more diverse waste types and environmental conditions and improving the quality of annotations could enhance performance. Integrating edge-computing solutions could enable real-time processing for dynamic applications.

9 Conclusion

This study presents a comprehensive approach to waste detection and classification using deep learning. While results indicate areas for improvement, the proposed system demonstrates the potential for scalable deployment in real-world waste management systems. Future work will focus on enhancing model robustness and expanding the system's applicability to broader contexts.

10 References

1. State-of-the-art computer vision model. YOLOv8. (n.d.). `https://yolov8.com/`
2. Resnet18 - Torchvision main documentation. (n.d.). `https://pytorch.org/vision/main/models/generated/torchvision.models.resnet18.html`
3. Trash annotations in context. tacodataset.org. (n.d.). `http://tacodataset.org/`