Smart Compost Bin for Food Waste Measurement

Aidan Beery, Patrick Donnelly

Oregon State University

Introduction

Food waste constitutes a major environmental and humanitarian crisis, threatening food security worldwide. An estimated 783 million people suffered from chronic hunger in 2022, while over 1 billion tons of food was wasted, generating over $1 trillion in economic losses. 60% of this waste is generated in the home, and yet there is no effective, scalable method for quantifying household food waste. Prior studies require users to maintain food waste journals, manually tracking each scrap of food. This process is tedious, expensive, and prone to user error. Meanwhile, a lack of high-quality food waste image datasets impedes the development of computer vision models to automate the task of food waste measurement.



Figure 1. The smart compost bin, a household food waste measurement device

We target three goals towards automatic household food waste measurement:

1. We design and manufacture the Smart Compost Bin, an AI-enabled device which automatically collects image and sensor data of disposed food.
2. To produce the first annotated, public dataset of post-consumer food images, we will conduct a pilot study with fifty households to annotate images of food waste using our smartphone app annotation platform.
3. We will use this dataset to train an open vocabulary semantic segmentation (OVSS) model to automatically measure household food waste.

We are recruiting participants for this study. If you are interested in volunteering to use the Smart Compost Bin, please scan this QR code to fill out our interest survey.

Hardware Design



(a) Raspberry Pi 5 (b) Camera (c) Gas Sensor (d) IR Camera (e) C2 Load Cell

We design a smart compost measurement device to facilitate the creation of a rich multimodal dataset on post-consumer food. Our device uses commercial-off-the-shelf sensors and 3D-printed components to reduce cost and streamline production. The base of the device contains a weight sensor, while cameras and sensors are embedded in the lid. A carbon air filtration system and dishwasher-safe removable liner improve sanitation, enhancing the user experience.

Data Collection

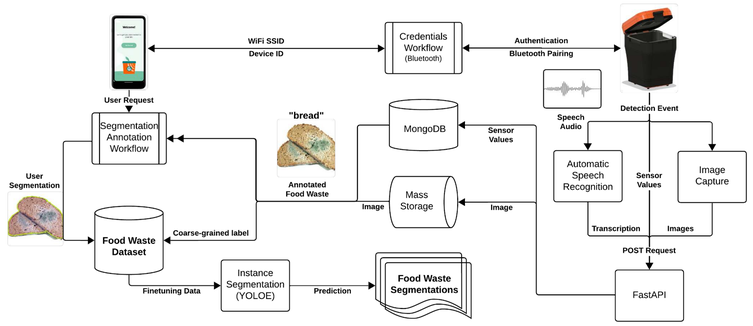


Figure 2. Data collection and annotation pipeline for images from compost bin

Each time a user deposits food, the bin captures a stereoscopic image pair, thermal images, and environmental measurements. The bin will prompt the user to describe the disposed item. An onboard microphone records the user's speech, which is transcribed using Whisper, an automatic speech recognition model.

Segmentation annotations are collected through our mobile app, *CompostSense*. This app tasks users to annotate their compost images by drawing an outline around each food item. A combination of coarse-grained open-vocabulary labels and detailed image segmentation annotations enables multi-stage fine-tuning of multimodal vision-language models for food waste recognition.

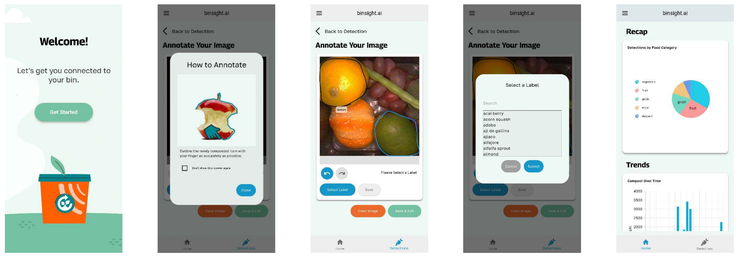


Figure 3. Companion mobile app for data annotation and device management

Food Waste Recognition

We utilize `YOLOE`, an open-vocabulary instance segmentation (OVIS) model, to classify compost items in commingled scenes. The food domain comprises a vast, complex label space. OVIS models leverage multimodal pretraining with natural language grounding to enable zero-shot recognition of unseen classes.

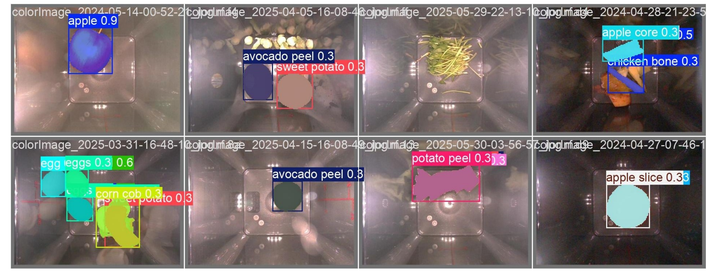


Figure 4. Example predictions from `YOLOE`, our open-vocabulary instance segmentation model

We evaluate the potential use of uneaten food image data to fine-tune `YOLOE` for compost image recognition tasks. YOLOE uses a small adapter network to transfer CLIP text embeddings to a joint region-object space, leveraging CLIP's multimodal image-text encoding capability to perform OVIS. We fine-tune `YOLOE` on three common food instance segmentation datasets with a total **47,329 instances** from **23,637 images** over **389 classes**. We fine-tune YOLO26x, a conventional CNN for instance segmentation, as a baseline, reporting these results in Table 1.

Dataset	Precision	Recall	mAP50	mAP50-95
Food201	0.478	0.349	0.347	0.255
UECFoodPix	0.803	0.612	0.673	0.582
FoodSeg103	0.576	0.410	0.429	0.352

Table 1. Food instance segmentation results of task-specific `YOLO26x-seg`, our baseline model

When compared to `YOLO26x`, `YOLOE` averages 12.1% lower mAP50 scores in food instance segmentation, indicating a performance trade-off when using open vocabulary models. When evaluating this model on *compost-test*, a dataset of 467 annotated images captured from the Smart Compost Bin, we find that `YOLOE` attains modest zero-shot performance on compost instance segmentation, but this performance collapses after fine-tuning on uneaten food datasets. This demonstrates the inability of food image recognition models to transfer to compost recognition tasks, highlighting the need for a robust dataset of annotated food waste images.

Dataset	Precision	Recall	mAP50	mAP50-95
Food201	0.349	0.329	0.283	0.206
UECFoodPix	0.766	0.561	0.623	0.530
FoodSeg103	0.494	0.395	0.384	0.308
compost-test	0.010	0.205	0.021	0.016
Zero-Shot (baseline `YOLOE`)				
compost-test	0.254	0.204	0.105	0.088

Table 2. Results of a `YOLOE` OVIS model fine-tuned for food recognition tasks

References

[1] Aidan J. Beery et al. Smart compost bin for measurement of consumer food waste. In *Companion Proceedings of the 26th International Conference on Multimodal Interaction*, ICMI Companion '24, page 100–107. Association for Computing Machinery, November 2024.