YOLO-recycler

Dominic Tran and Steven Siadaty
Monday, December 15th

Motivation

- People may not always know when an item is able to be recycled or which category to recycle it under (paper, metal, plastic, etc)
- When recycling, one incorrect item in a batch (ie metal in a paper recycling batch) ruins the entire batch
- With single-stream recycling, people must manually sort it
- So, we are trying to find a way to make it easier for people to sort the items before they put it in the bin
- Contamination rates exceed 25% in most U.S. recycling facilities.
- The result: billions in lost recyclable materials, more landfill waste, higher carbon cost.

Related Work

- StreetView-Waste: A Multi-Task Dataset for Urban Waste Management
  - https://arxiv.org/html/2511.16440v1
- Recyclable waste image recognition based on deep learning
  - https://www.sciencedirect.com/science/article/abs/pii/S0921344921002457
- Real-Time Household Waste Detection and Classification for Sustainable Recycling: A Deep Learning Approach
  - https://www.mdpi.com/2071-1050/17/5/1902
- YOLOv8-based Waste Detection System for Recycling Plants: A Deep Learning Approach
  - https://www.researchgate.net/publication/374919875_YOLOv8-Based_Waste_Detection_System_for_Recycling_Plants_A_Deep_Learning_Approach

Background

- Instance segmentation assigns a class label and pixel-level mask to each object
- YOLOv8-seg performs real-time instance segmentation using a single-stage architecture
- Waste datasets contain complex scenes with overlapping objects
- Performance is typically evaluated using mAP (mean Average Precision) and IoU (Intersection over Union) metrics

Claim / Target Task

- Claim: After task-specific fine-tuning, the model can reliably distinguish recyclable from non-recyclable objects in real-world scenes with occlusion and background clutter.
- Target Task: Instance segmentation of recyclable objects in complex real-world scenes, where the model takes RGB images as input and outputs pixel-level masks and class labels for recyclable and non-recyclable objects.

Which Items are Recyclable?

Proposed Solution

Fine-tune a state of the art segmentation model on synthetically-compiled data, for detecting recyclable objects, making recycling easier and more accurate

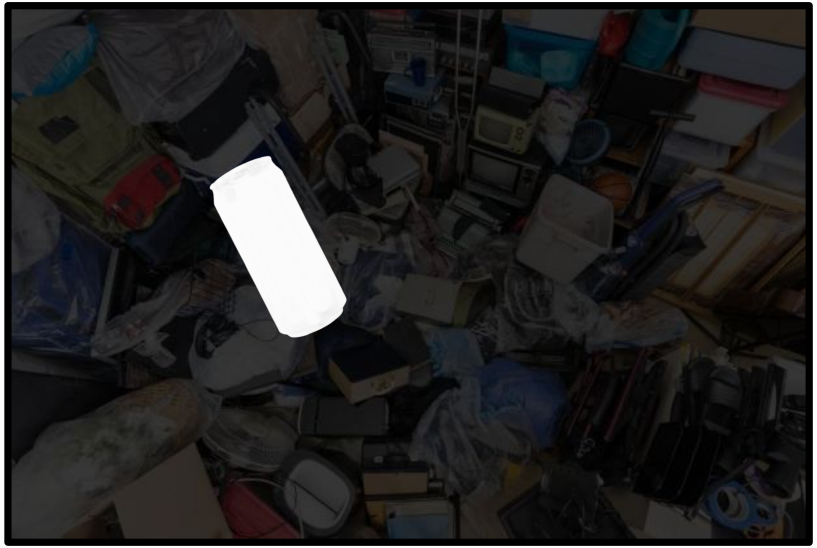
Implementation

- Start with the TrashNet dataset
  - Clean and relabel
  - Generate segment-masks with GrabCut
  - Create synthetic training data; superimpose isolated masks onto noisy backgrounds
- Fine-Tune Classification Model
  - Use YOLOv8-Seg pretrained weights.
  - Train for accurate segmentation + confidence calibration.
- Real-World Testing
  - Test on real-world images (non-synthetic)
  - Measure accuracy, speed, and confidence in real waste scenarios.

Data Synthesis Pipeline

- Foreground extraction:
  - Run background removal with BRIA RMBG-2.0 model to produce foreground-alpha masks and RGBA cutouts for each plastic image
- Synthetic compositing:
  - For each plastic image, generate 5 composites on random backgrounds
  - Augmentations: scale/rotation/position/color jitter
- Split data into train/test/val
- Generate mask-labels to prep for YOLO training

Synthetic Data Example



Data Summary

Example Trashnet Image:

Trashnet Composition:

- cardboard: 806
- glass: 1002
- metal: 820
- paper: 1188
- **plastic: 964**
- trash: 274

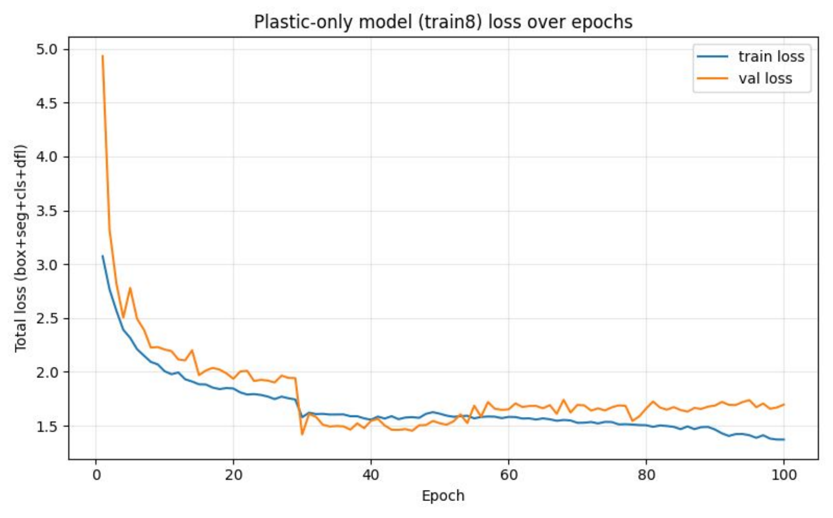
Plastic selected for this proof of concept because the images were consistently in-frame, and plastic recyclable objects are very commonly found in the real world

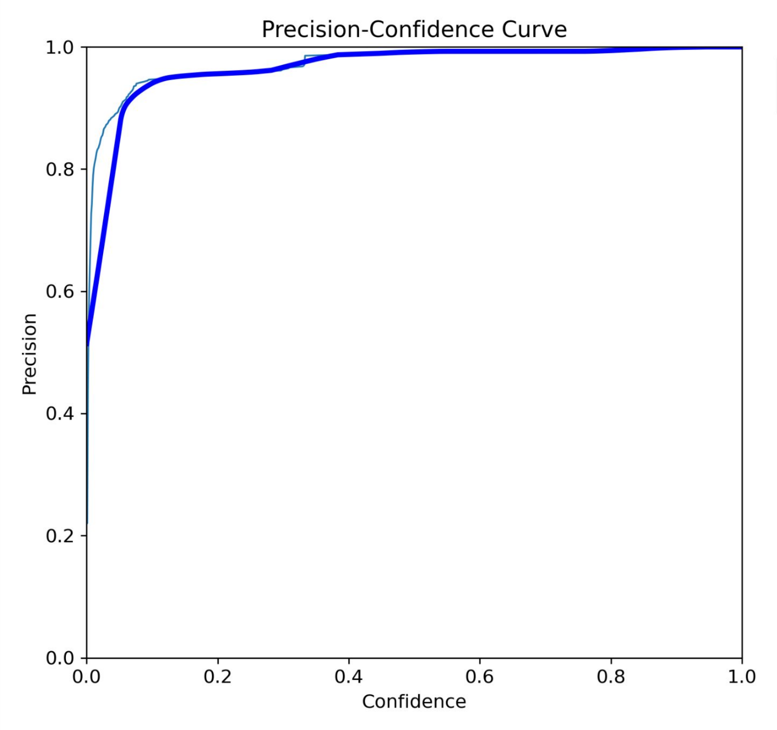
964 plastic images + 24 background images → 4820-image synthetic dataset

Total Images per Split:

- Train - 3860
- Val -420
- Test - 540

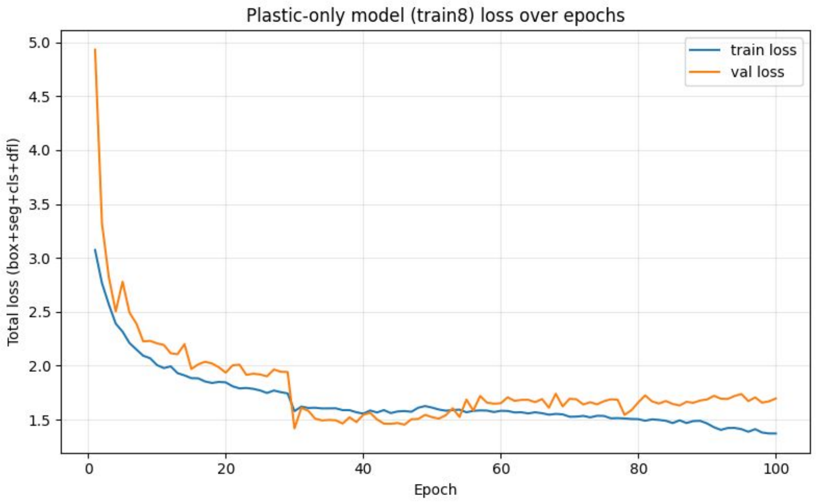
Experimental Results





Experimental Analysis

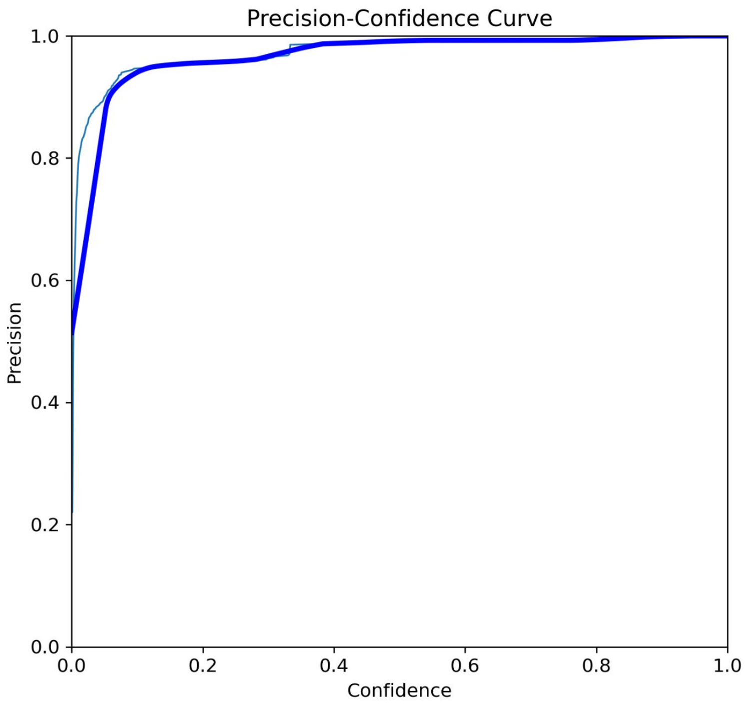
- Best epoch: 30
  - Train→ 1.5782
  - Val: → 1.4177
- box_loss: Localization loss for bounding boxes
- Seg_loss: Per pixel segmentation mask loss
- cls_loss: Classification loss over the class logits for each detected instance (background vs plastic)
- dfl_loss: Location prediction as combination of discrete classifications



Experimental Analysis

- Proportion of correct predictions given confidence threshold
- Precision converges very quickly
- Confidence threshold ~0.8 selected
    - ~.99 Precision

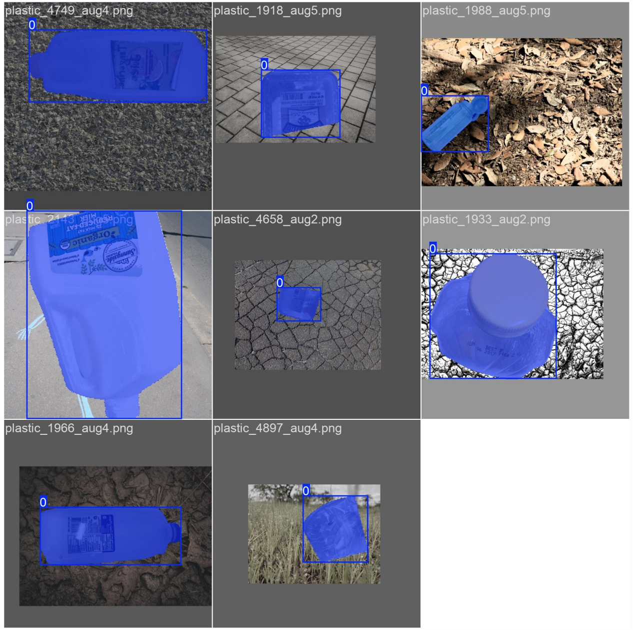
- metrics/recall(B|M)=0.25195
- metrics/precision(B|M)=0.99971



Test Run on Metal/Glass Synthetic Data

- Plastic-only model was testing on dataset of metal and glass synthetic images
- ~99% of images had objects falsely-identified as plastic objects
  - Extreme false-positive issue
  - Poor negative-case performance
- Model appears to learn to select objects from the background rather than plastic-objects in particular
- Multi-class training will potentially improve cross-class performance
- Recall value ~.252, also indicating poor false-positive performance

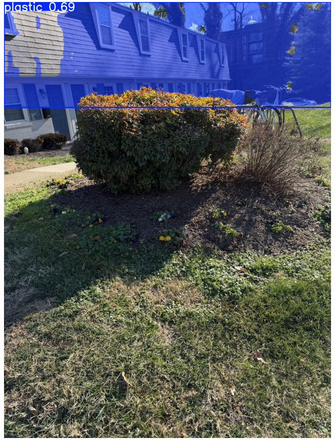
Example Output



Synthetic

Real-World

More Examples



Conclusion and Future Work

- Model was reasonably accurate at selecting objects in real-world scenes, but failed to classify between plastic objects and other materials
- Currently, we only used synthetic images (background with item mask), so we'd like to expand that to real images of recyclable objects on a complex background
- We also only trained on plastic masks, so we'd like to expand that to paper/cardboard and metal recyclables as well, which will likely fix the classification issue

References

- Trashnet: github.com/garythung/trashnet?tab=readme-ov-file
- YOLOv8: https://github.com/ultralytics/ultralytics
- RMBG-2.0: https://huggingface.co/briaai/RMBG-2.0

Video Demo

- Slide Deck Walkthrough
  - https://youtu.be/YcraLptXeWg
- Code Walkthrough
  - https://youtu.be/L-Bif0_NOkU

Work Distribution

- Dominic
  - YOLOv8-seg setup, fine-tuning, and evaluation
- Steven
  - Synthetic dataset creation with instance-level masks