Instance Segmentation for Campus Waste Detection Using Mask R-CNN

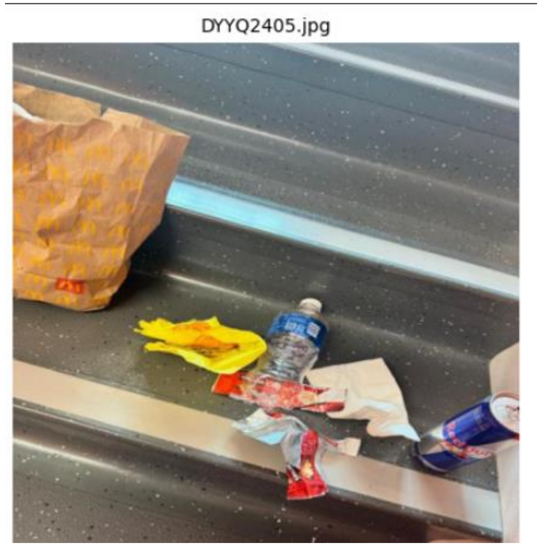
By

Yeswanth Labba

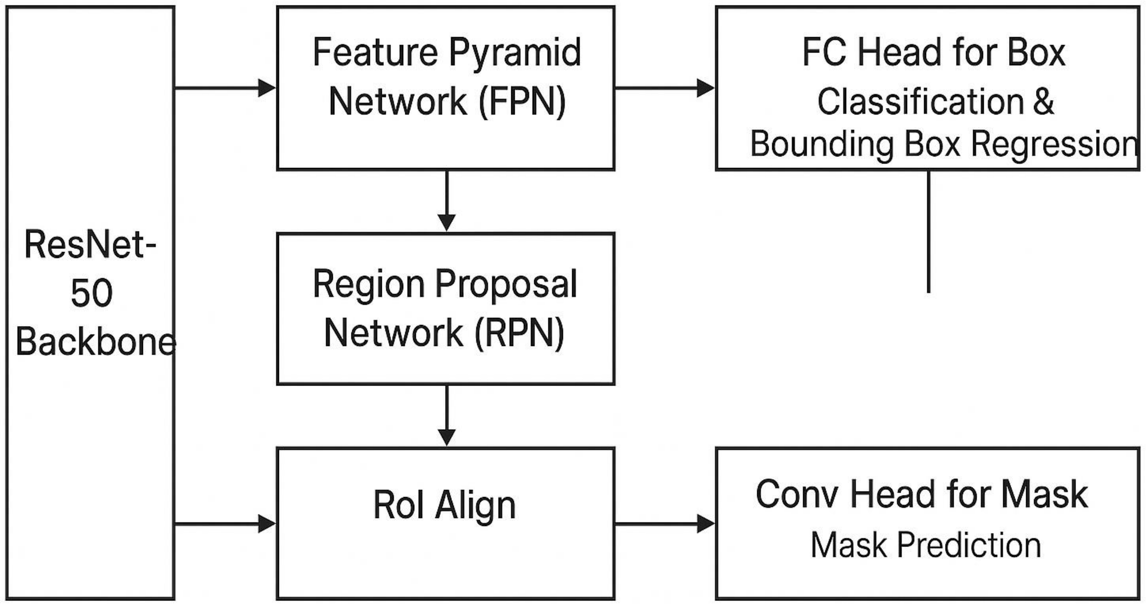
Nagasai Jaja

Dataset Overview

- **Task:** Detect and segment waste items on UNH campus images
- **Data source:** COCO-format annotations (remapped from original LabelMe)
- **Number of samples:** 200 images (512×512 px each)
- **Classes:** 2 (recyclable_container, food_packaging)
- **Partition:**
    - Train: 160 (80 %)
    - Val: 20 (10 %)
    - Test: 20 (10 %)
- **Normalization:** Images normalized using ImageNet mean [0.485, 0.456, 0.406] and std [0.229, 0.224, 0.225] (Detectron2 defaults)
- **Augmentation (training):** Random horizontal flips; resize shortest edge to 800 px; other default FPN augmentations from Mask R-CNN config



Model Architecture

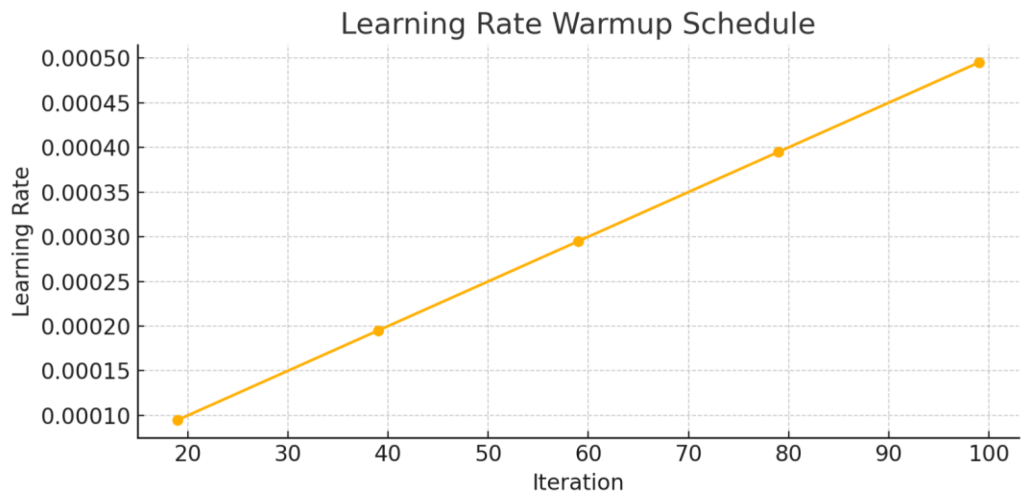


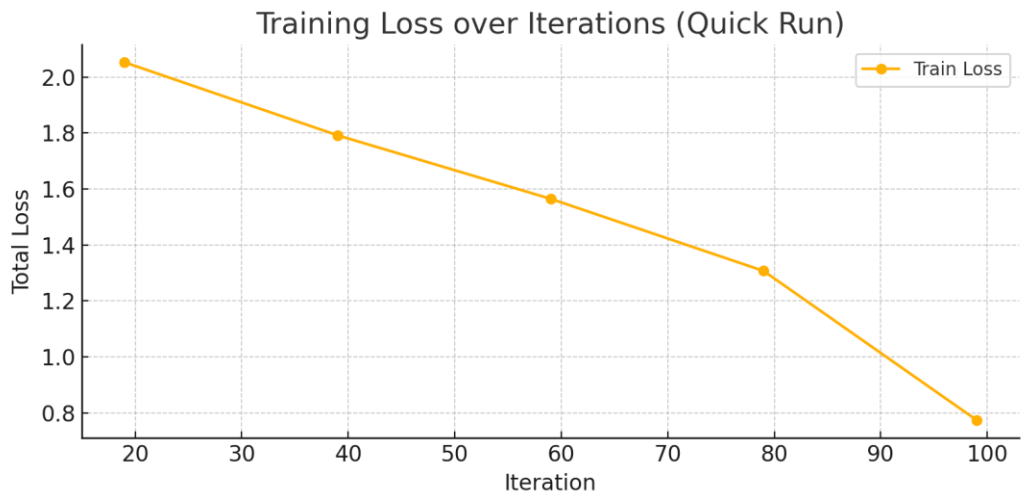
Mask-R CNN

- **Base network:** Mask R-CNN with ResNet-50 backbone + Feature Pyramid Network (FPN)
- **Pretrained weights:** COCO-InstanceSegmentation/mask_rcnn_R_50_FPN_3x
- **ROI Heads:**
- Box head → classification + bbox regression
- Mask head → per-pixel segmentation
- **Num classes:** 2
- **Losses:** Cross-entropy for class, smooth L1 for boxes, binary cross-entropy for masks

Training Setup

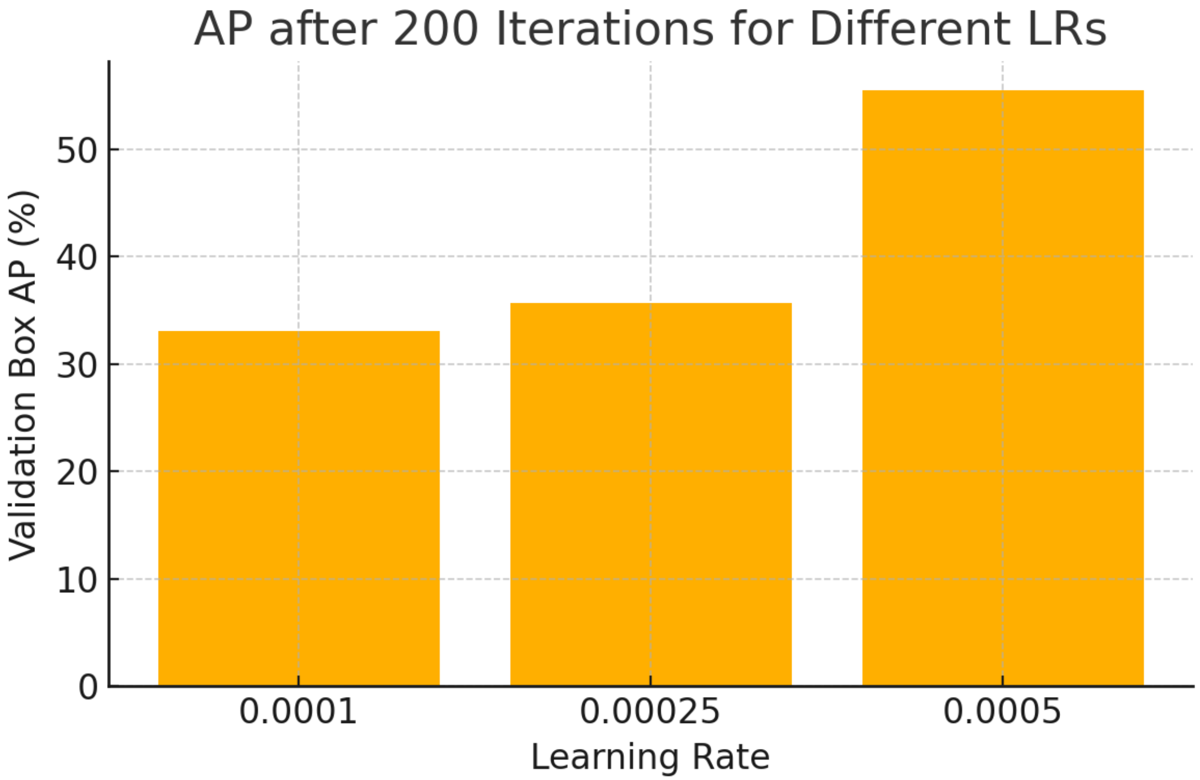
- **Framework:** Detectron2 DefaultTrainer
- Initial hyperparameter grid search over learning rates [1e-4, 2.5e-4, 5e-4]
- Short runs: batch size 4 images/GPU for 200 iterations
- Selected best LR (5×10^{-4}), then merged train+val (180 images) for fine-tuning
- **Final run:** 1,000 iterations, SGD optimizer (momentum 0.9, weight decay 1×10^{-4})





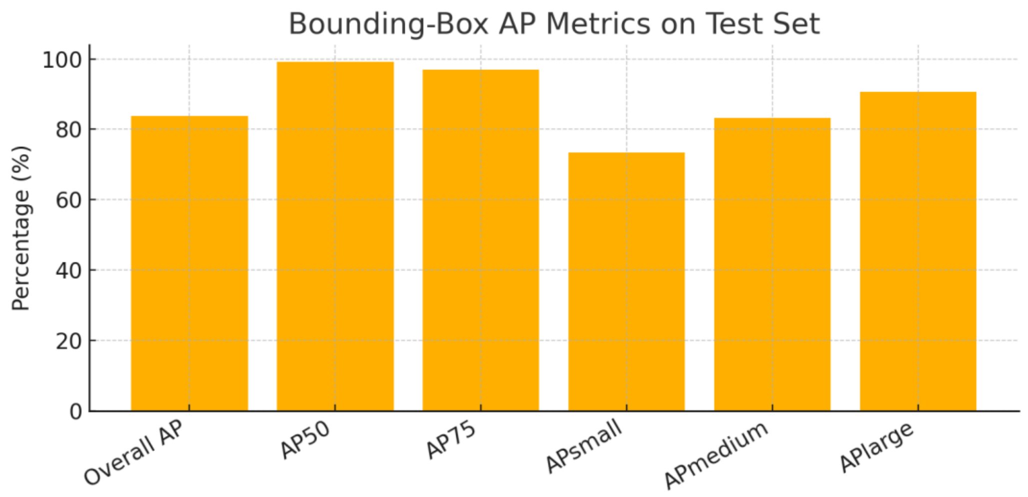
Hyperparameter Tuning Results

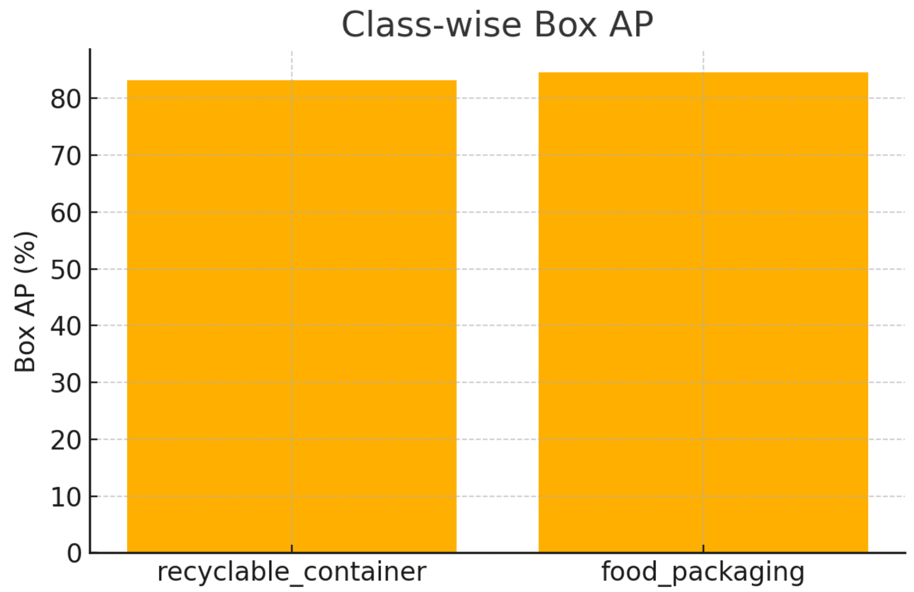
- LR = 1×10^{-4} → box AP ≈ 33 % after 200 iterations
- LR = 2.5×10^{-4} → box AP ≈ 36 % after 200 iterations
- LR = 5×10^{-4} → box AP ≈ 55 % after 200 iterations
- Clear performance boost at higher learning rate, indicating under-learning at lower rates



Final Test-Set Performance (Bounding Boxes)

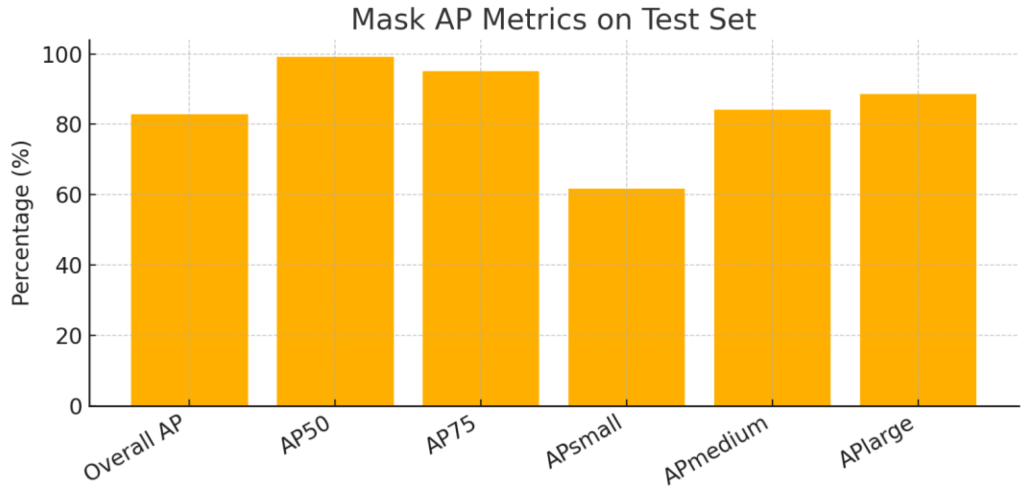
- Overall AP: 83.9 %
- AP50: 99.1 % | AP75: 97.1 %
- APsmall/APmedium/APlarge: 73.4 % / 83.3 % / 90.7 %
- Class-wise box AP:
    - recyclable_container: 83.2 %
    - food_packaging: 84.5 %

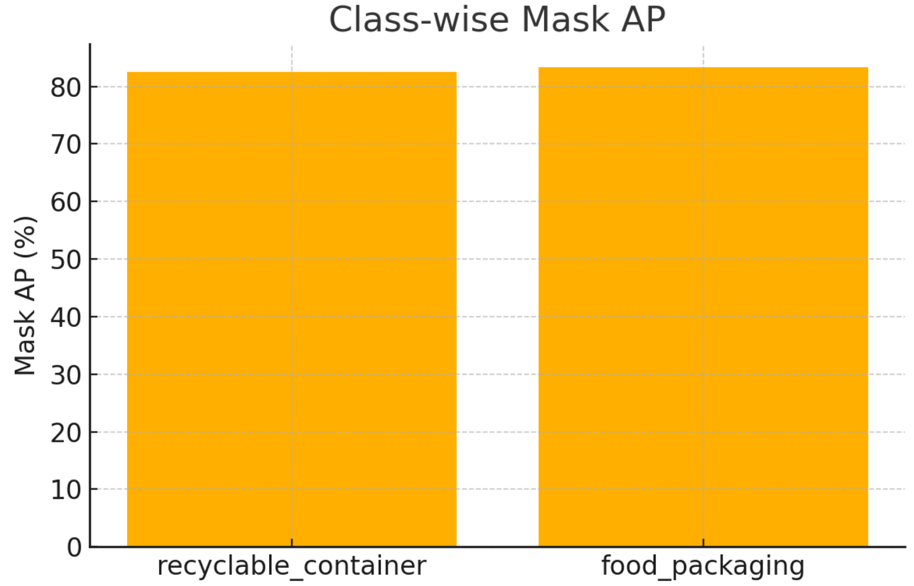




Final Test-Set Performance (Segmentation Masks)

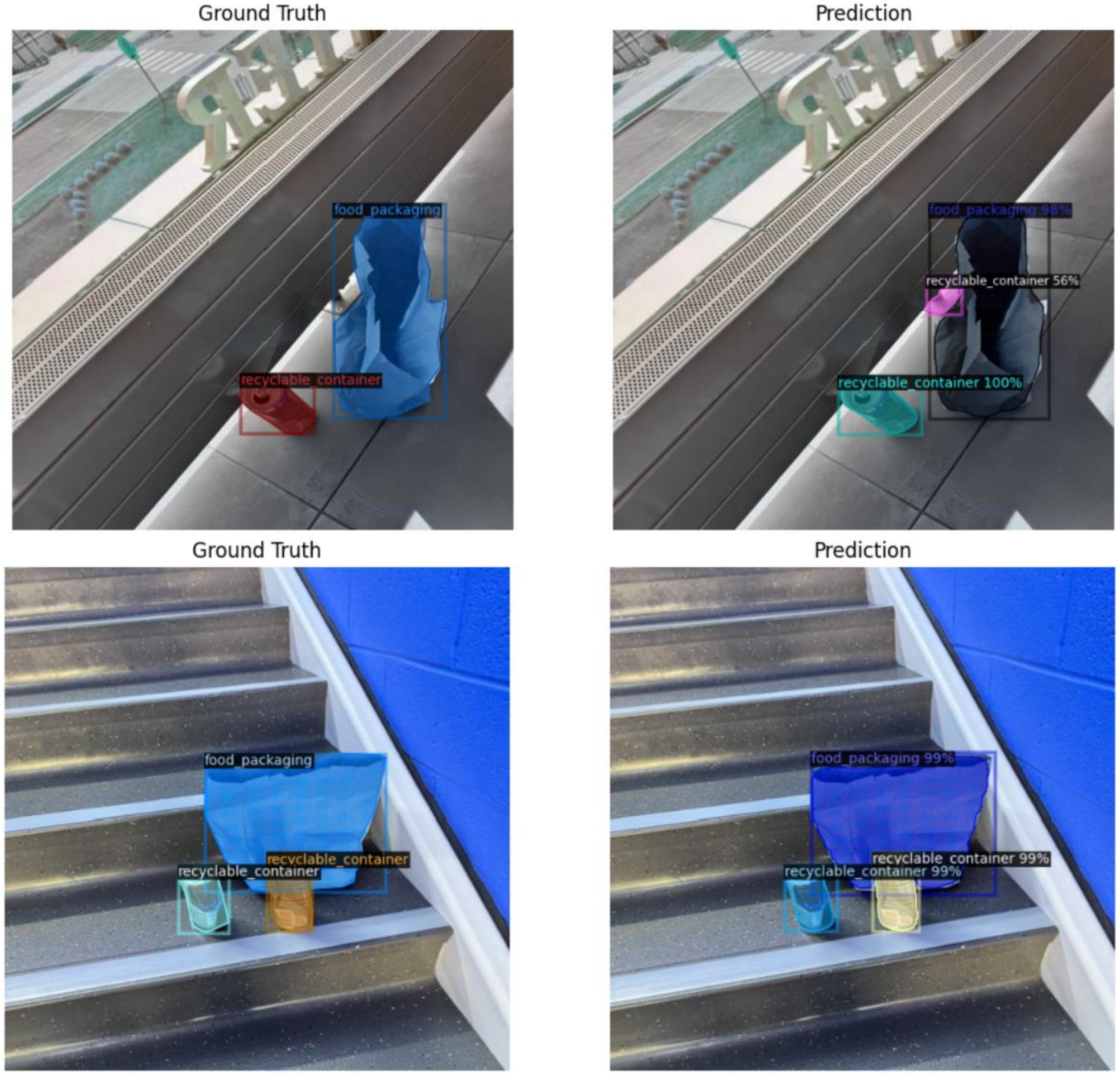
- Overall mask AP: 82.9 %
- AP50: 99.1 % | AP75: 95.2 %
- APsmall/APmedium/APlarge: 61.8 % / 84.1 % / 88.6 %
- Class-wise mask AP:
    - recyclable_container: 82.6 %
    - food_packaging: 83.3 %





Qualitative Results

- Display side-by-side: ground-truth vs. predicted masks
- High-quality boundary delineation even under occlusion and varied lighting
- Rare misclassifications; model reliably ignores background clutter



Conclusions & Future Work

- Demonstrated effective transfer learning with just 200 annotated images
- Achieved > 80 % AP on both detection and segmentation tasks
- Future plans:
  - Expand to additional waste categories (paper, metal, etc.)
  - Deploy optimized model on edge devices for real-time campus monitoring
  - Integrate alerts into campus recycling management system