16/01/2024

Vision and Perception Final Project

WasteWatch: an Algorithm for Waste Detection and Classification

SAPIENZA
UNIVERSITÀ DI ROMA

Students: Alessia Ciarlo - 1796690
Alessandro Garbetta- 1785139

Contents:

- Aim of our project
- Datasets
- First step: Detection Model
- Second step: Classification Model
- Final application part
- Results and Conclusions

The aim of our project

In recent years, waste levels have been increasing rapidly, and the only solution is an adequate level of recycling, a process that is ***not yet fully carried out autonomously***.

This led us to our project idea: implementing a <u>model that performs the detection and classification of waste</u>, categorizing it based on its construction material, which can then be used during the sorting of recyclable materials.

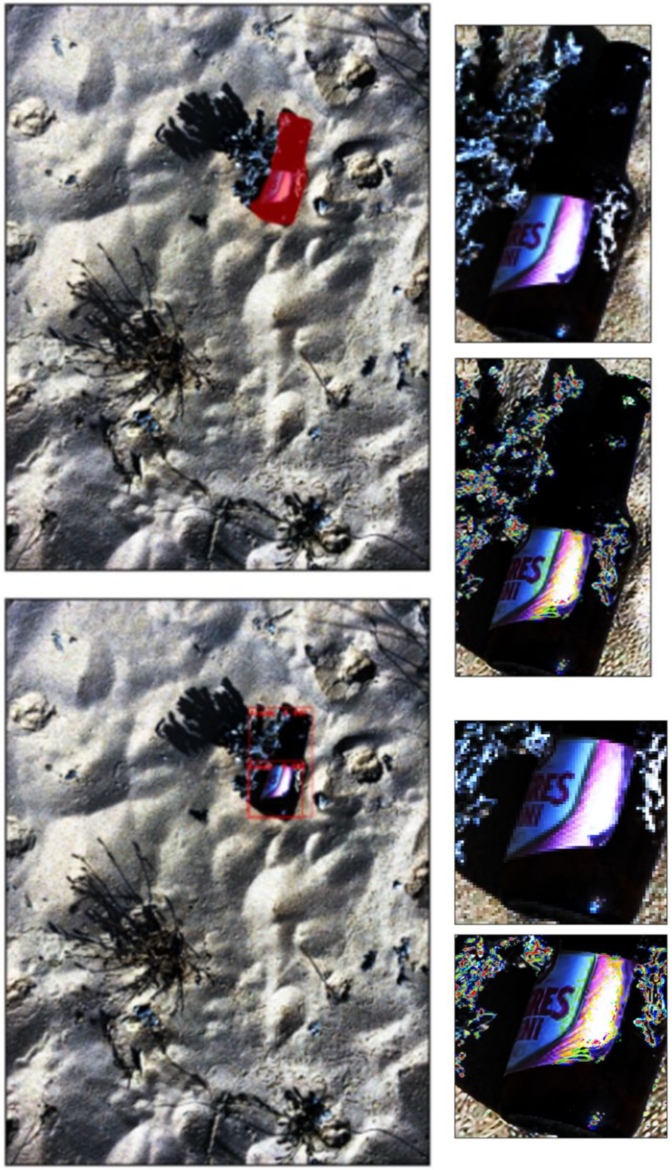
Our primary goal is to be able to recognize and differentiate objects in outdoor environments (fields, forests, sea, sand, etc.). Hence, the choice of datasets to use has been crucial.

Our ideas include:

- ***2 step algorithm approach***: detection part and classification part divided (and then unified in the WasteWatch application)

- working on objects composed of multiple parts or multiple materials to ***recognize individual components***;

- work with images of small-sized waste and apply ***super resolution techniques*** to identify and classify them.



Data

Several datasets were considered, both for the detection and segmentation part, and for the classification part.

- TACO
- MJU
- Drinking Waste
- Trash-Net
- Waste Pictures
- UAVVaste
- Trash-Can
- Trash-Icra

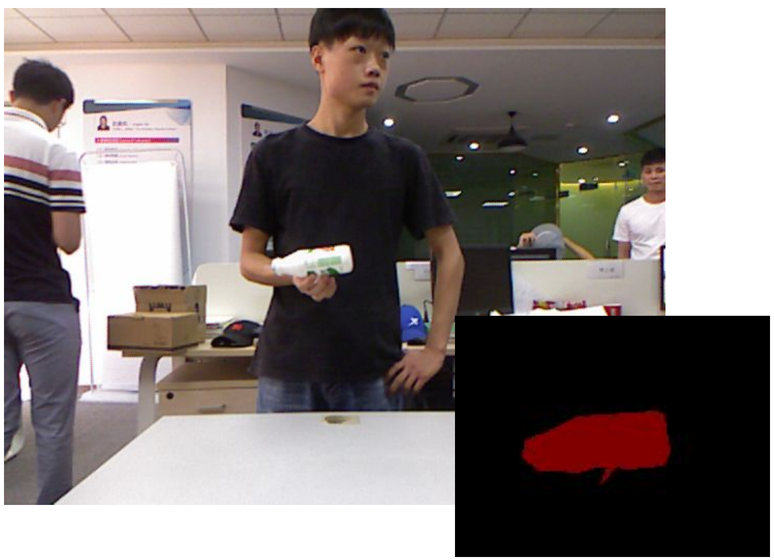
Detection Datasets

- TACO
- MJU

Outdoor

Indoor

Both with annotations in COCO format



Classification Datasets

- Trash-Net
- Waste-Pictures
- Drinking-Waste

- Mix-Waste-Pictures

- TACO

some of dataset contains different “version” of the same images

Classification labels

Mix-Waste-Pictures → 6

1. Biodegradable
2. Glass
3. Paper
4. Plastic (bag, bottle, metal)
5. Special
6. Trash

datasets with different original classes. Reorder and distribution

TACO → 4/6

1. Glass
2. Paper
3. Plastic (bag, bottle, metal)
4. Trash

Dataset Unbalanced, resolved in part with csv distribution classes and oversampling method

Trash-Net → 6

1. Cardboard
2. Glass
3. Metal
4. Paper
5. Plastic
6. Trash

Object detection part
(for waste detection)

Problems:

- several objects in the environment
- deformed objects
- composed objects (milk cartoon)

Dataset used:

- TACO
- MJU

Solution:

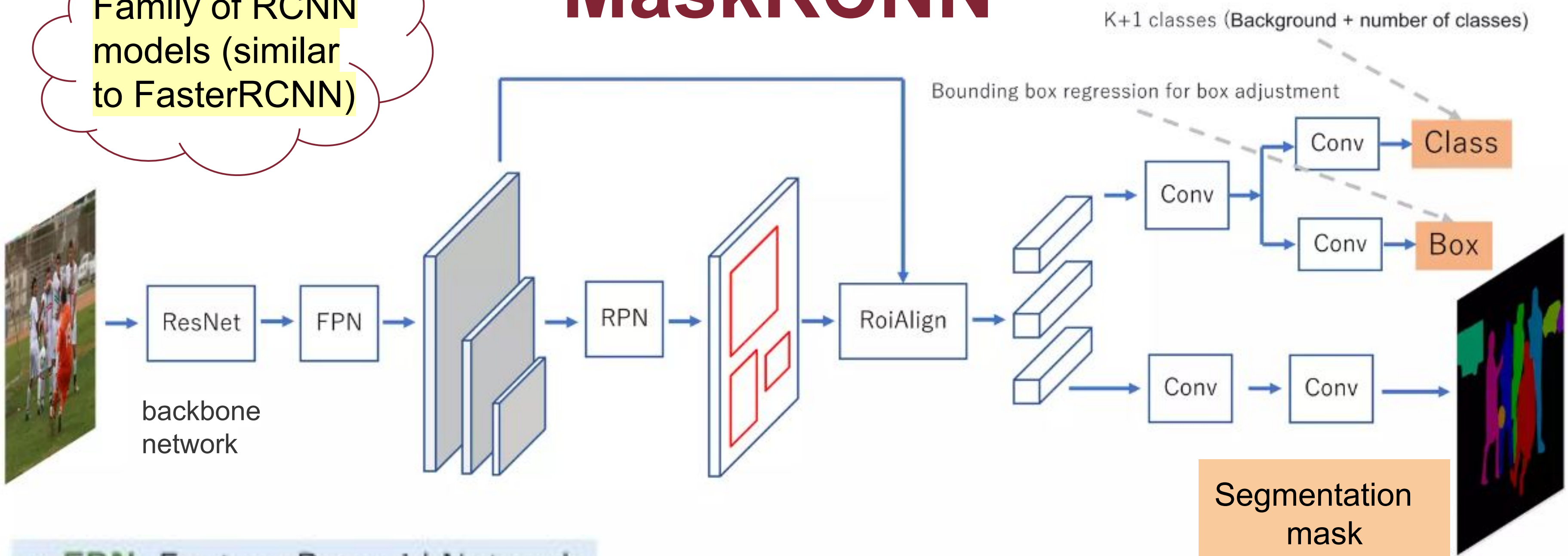
use a model that allow us to have both:

- position of the object (bounding box)
- specific shape of the object (instance segmentation)

MaskRCNN

MaskRCNN

Family of RCNN models (similar to FasterRCNN)



- **FPN**: Feature Pyramid Network
- **RPN**: Region Proposal Network
- **RoI**: Region of Interest

It creates a multi-scale feature pyramid by combining features from different levels of the ResNet50 network

It generates region proposals that might contain objects within the image. It operates on the feature map coming from the FPN.

The RoIAlign aligns the features within a region of interest (RoI) with the spatial grid of the feature map. This prevents information loss that can occur.

Data preparation and training

TACO

- Waste-in-environment images resized to 640x480 (like the MJU images)
- Creation of ***Mask images*** (640x480) starting from segmentations in COCO format
- Preprocessing of these mask images to create tensors of binary masks

MJU

- Preprocessing of mask images, already present in this dataset, to create tensors of binary masks (N masks = N objects)

Input: *(img, bbox, masks, label)*

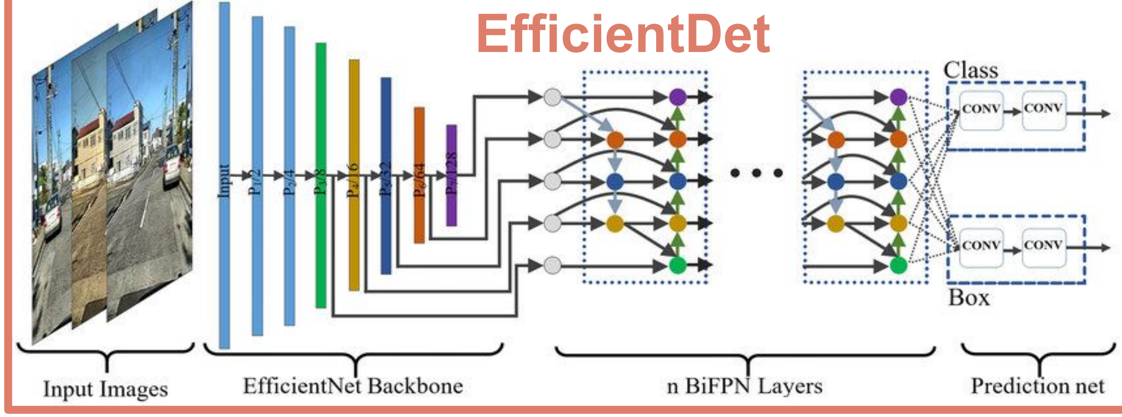
→

- *5/10 epochs*
- *4 batch (memory problems)*
- *SGD optimizer (momentum)*
- *mAP calculation*

→

Output: *(boxes, labels, scores, masks)*

Results
Detection Part



Evaluation using:
IoU - Intersect over Union
mAP - mean Averaged Precision

	MaskRCNN (5epochs)	**MaskRCNN** (10epochs)	**EfficientDet** (5epochs)
TACO	0.97	0.984	0.91
MJU	0.94	0.973 (with overfitting)	—

The best performances, also testing on images from other datasets, are achieved by the training on TACO (10epochs).

We have tried to train on a mix of the 2 datasets
→ problems of computational resources and differences between the 2 datasets.

Classification part

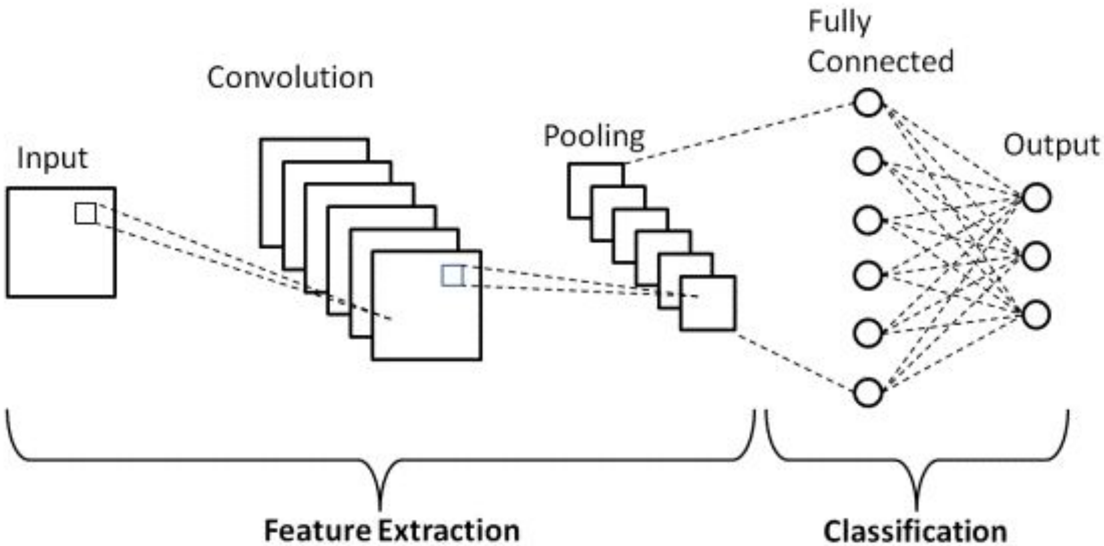
CNN architecture model

Problems:

- type of data
- diversity
- features of image (such as brightness , …)

Solution:

- use dataset with different images and orientations
- balanced classes
- normalization data techniques

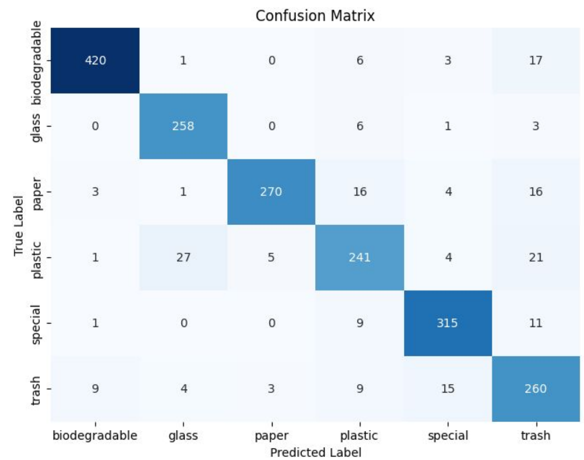


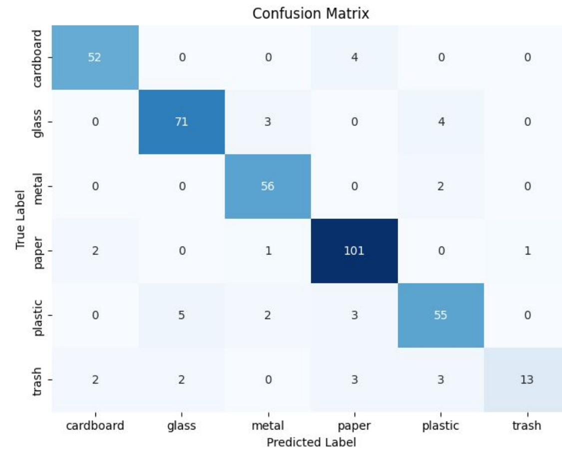
3 CNN architecture models were examined: ResNet50, VGG16 and EfficientNet.
Each image underwent a resizing of 224 by 224, and normalization of the three RGB channels was applied.

We use CrossEntropyLoss function, SGD optimizer and ReduceLROnPlateau to optimize the learning rate’s scheduler.

Results
Classification Part

Evaluation using:
Accuracy
F1-Score



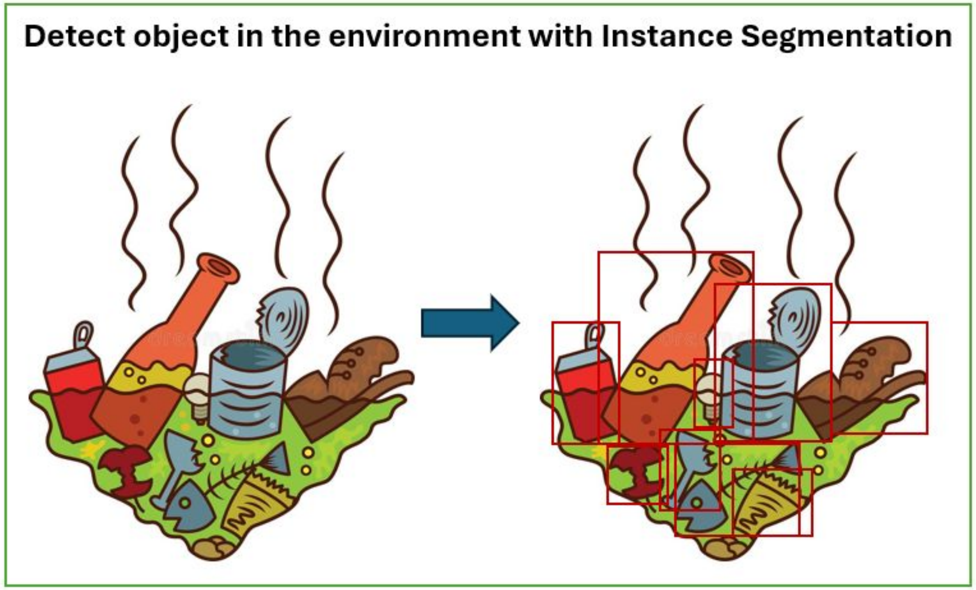


	RESNET 50	VGG16	EFFICIENT NET
MIX-WASTE-PICTURES	0,92	0,91	0,89
MIX-WASTE-PICTURES-CROP	0,82	0,82	
TACO	0,78		
TACO4	0,82		
TRASH-NET	0,9	0,85	

WasteWatch
(2 steps approach)

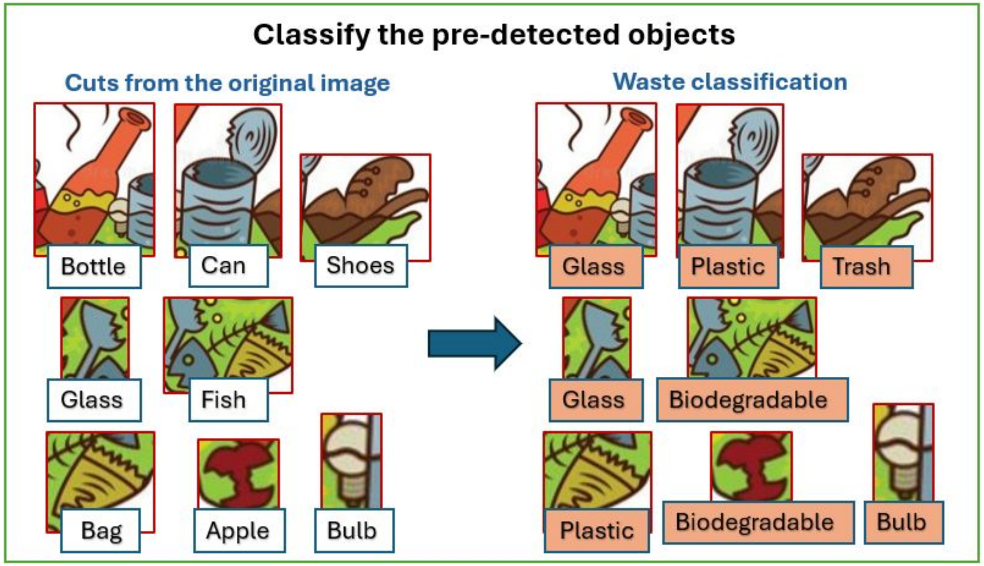
- Models chosen for the final part
- Overall results
- Some final examples

Step 1: Object Detection



In the **1° step**, the image passes through the Object Detector (MaskRCNN trained model) and is cutted using the bounding boxes provided in output.

Step 2: Object Classification

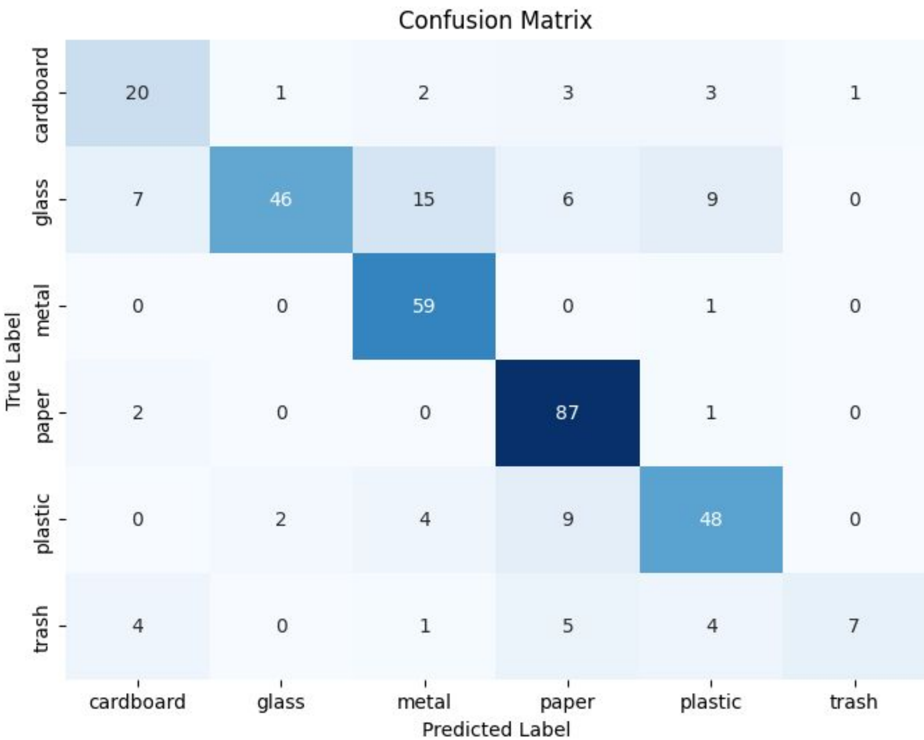


In the **2° step**, the cutted images pass through the Classifier (VGG16 trained model) and we have in output the corresponding types of waste for each of them.

Overall Results

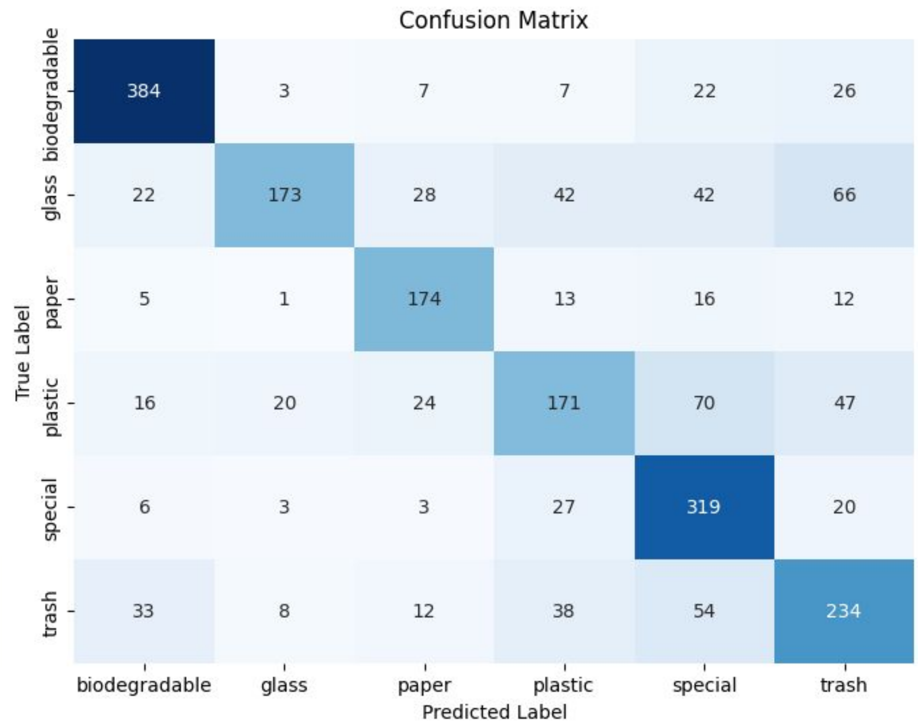
Dataset	Pipeline	Result
MIX WASTE PICTURES	MaskRCNN_TACO + VGG_MIX_BAL + no super resolution	63%
	MaskRCNN_TACO + VGG_MIX_BAL + super resolution	70%
TACO	MaskRCNN_TACO + RES_cropTACO_over + super resolution	63%
	MaskRCNN_TACO + RES_cropTACO4 + super resolution	67%
THRASH NET	MaskRCNN_TACO + RES_THRASH_NET + super resolution	77%
	MaskRCNN_TACO + VGG_THRASH_NET + super resolution	68%

Results



MaskRCNN_TACO + RES_THRASH_NET

+ super resolution (77%)



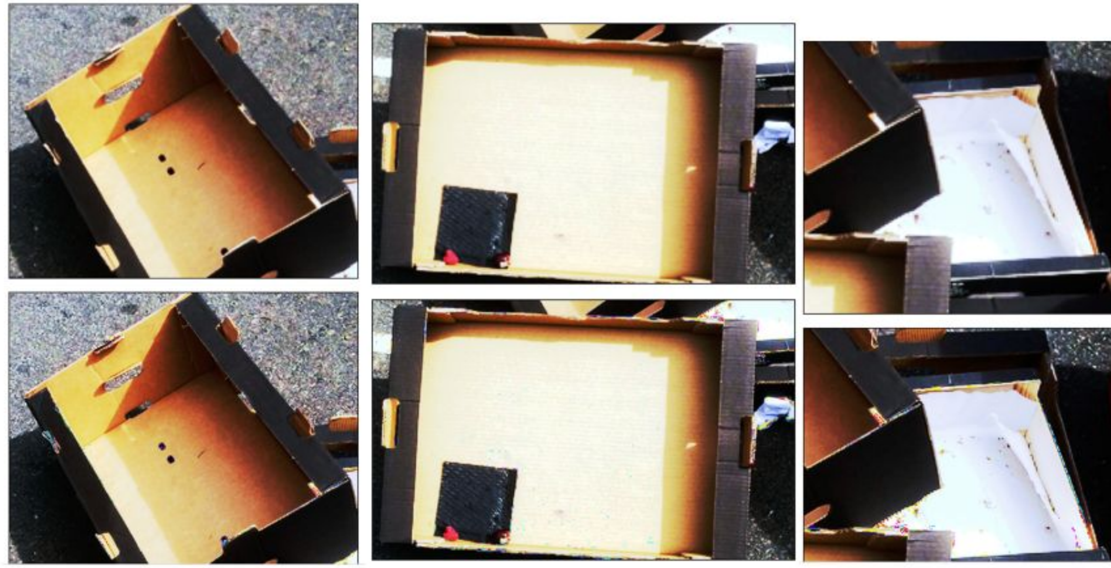
MaskRCNN_TACO + VGG_MIX_BAL

+ super resolution (70%)

Example 1:

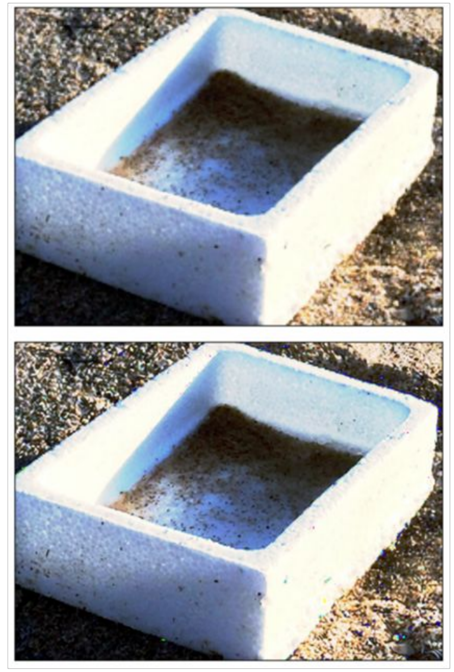
```
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_1/000001.jpg
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_1/000001.jpg
THE NET PREDICTS special
```

Example 2:



```
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_1/000102.JPG
THE NET PREDICTS paper
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_1/000102.JPG
THE NET PREDICTS paper
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_1/000102.JPG
THE NET PREDICTS paper
```

Example 3:

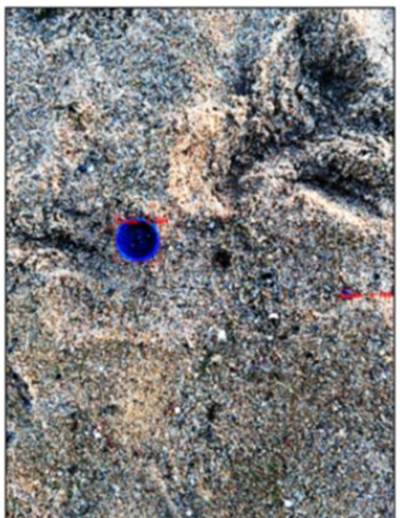


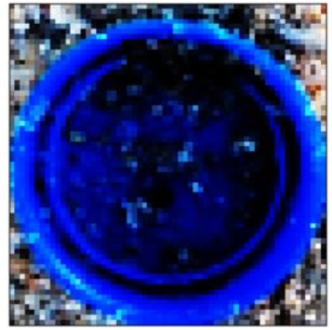
```
IDENTIFIED OBJECTS:  1
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/
THE NET PREDICTS plastic
```

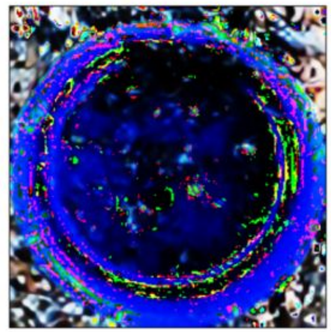
Example 4:

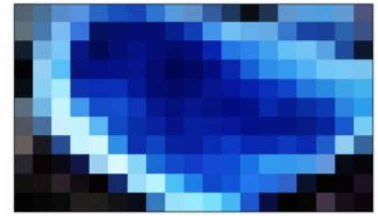
```
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000022.JPG
THE NET PREDICTS trash
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000022.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000022.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000022.JPG
THE NET PREDICTS paper
```

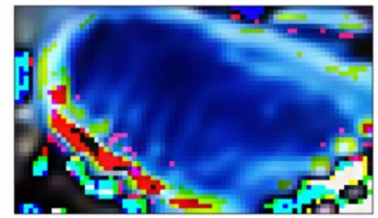
Example 5:





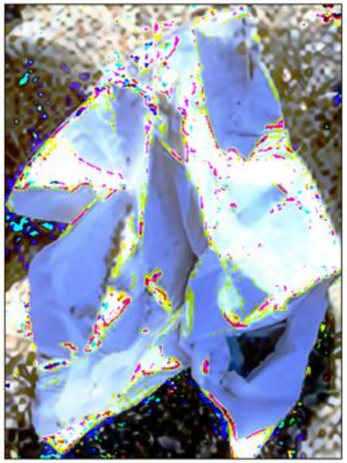






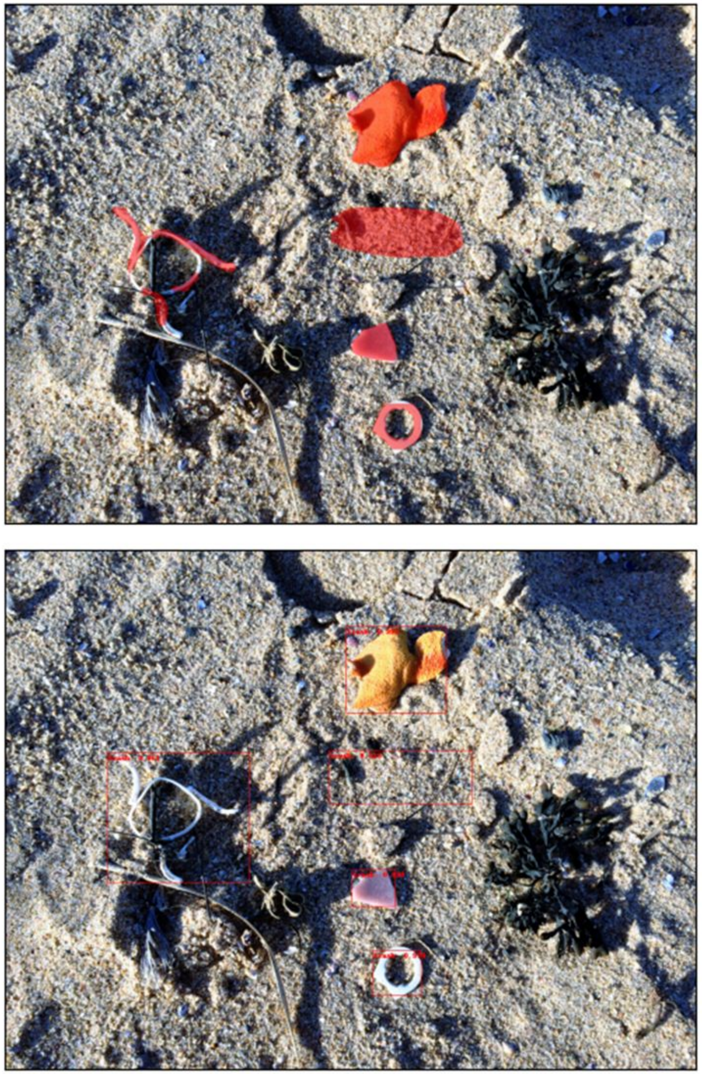
```
IDENTIFIED OBJECTS:  2
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000013.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000013.JPG
THE NET PREDICTS trash
```

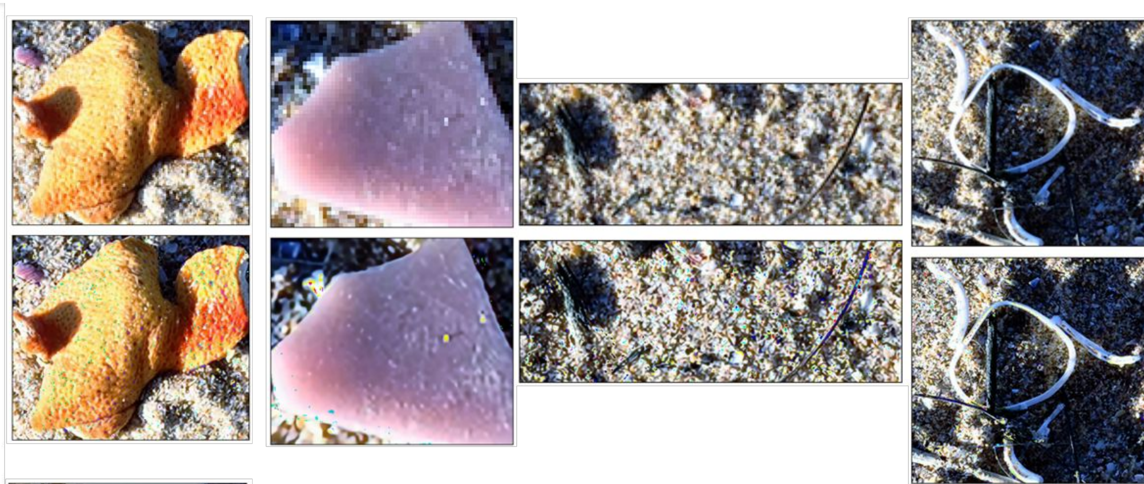
Example 6:

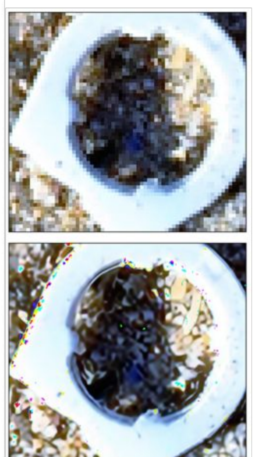


```
IDENTIFIED OBJECTS:  1
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/
THE NET PREDICTS trash
```

Example 7:





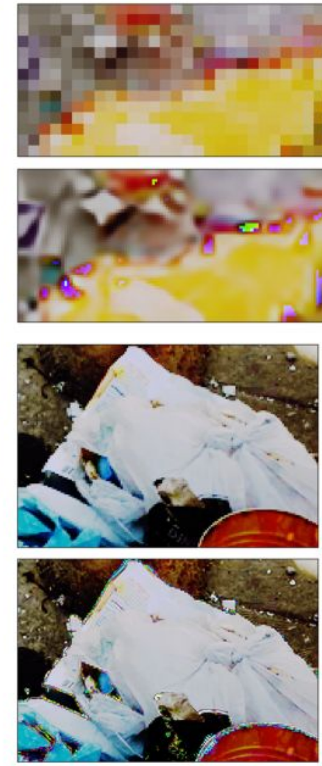


```
IDENTIFIED OBJECTS:  5
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000054.JPG
THE NET PREDICTS biodegradable
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000054.JPG
THE NET PREDICTS trash
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000054.JPG
THE NET PREDICTS biodegradable
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000054.JPG
THE NET PREDICTS trash
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_4/000054.JPG
THE NET PREDICTS plastic
```

Example 8:

```
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000012.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000012.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000012.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000012.JPG
THE NET PREDICTS plastic
```

Example 9:

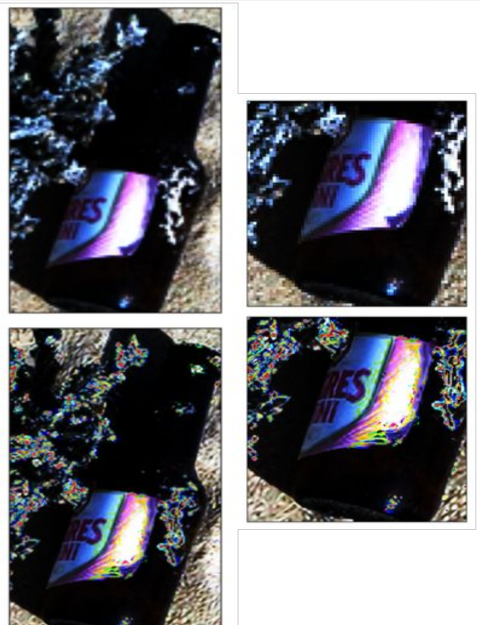


```
IDENTIFIED OBJECTS:  5
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000036.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000036.JPG
THE NET PREDICTS paper
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000036.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000036.JPG
THE NET PREDICTS biodegradable
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000036.JPG
THE NET PREDICTS plastic
```

Example 10:

```
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS paper
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_7/000078.JPG
THE NET PREDICTS paper
```

Example 11:

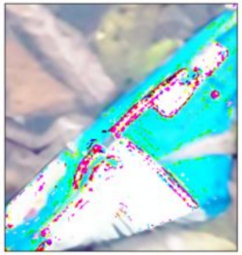


```
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000101.JPG
THE NET PREDICTS plastic
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_5/000101.JPG
THE NET PREDICTS plastic
```

Example 12:

```
IDENTIFIED OBJECTS:  2
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_3/IMG_4963.JPG
THE NET PREDICTS glass
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_3/IMG_4963.JPG
THE NET PREDICTS plastic
```

Example 13:



```
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_6/000029.JPG
THE NET PREDICTS paper
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_6/000029.JPG
THE NET PREDICTS trash
/content/drive/MyDrive/VISIOPE_Project/DATASET/TACO/batch_6/000029.JPG
THE NET PREDICTS trash
```

References

1. *Majchrowska, Sylwia & Mikołajczyk-Bareła, Agnieszka & Ferlin, Maria & Klawikowska, Zuzanna & Plantykow, Marta & Kwasigroch, Arkadiusz & Majek, Karol. (2022). Deep learning-based waste detection in natural and urban environments. Waste Management. 138. 274-284. 10.1016/j.wasman.2021.12.001.*

2. *Erkınay Özdemir, Merve & Ali, Zaara & Subeshan, Balakrishnan & Asmatulu, Eylem. (2021). Applying machine learning approach in recycling. Journal of Material Cycles and Waste Management. 23. 10.1007/s10163-021-01182-y.*

3. *R. Sultana, R. D. Adams, Y. Yan, P. M. Yanik and M. L. Tanaka, "Trash and Recycled Material Identification using Convolutional Neural Networks (CNN)," 2020 SoutheastCon, Raleigh, NC, USA, 2020, pp. 1-8, doi: 10.1109/SoutheastCon44009.2020.9249739.*

4. *M. Fulton, J. Hong, M. J. Islam and J. Sattar, "Robotic Detection of Marine Litter Using Deep Visual Detection Models," 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 2019, pp. 5752-5758, doi: 10.1109/ICRA.2019.8793975.*

5. *Hong, Jungseok & Fulton, Michael & Sattar, Junaed. (2020). TrashCan: A Semantically-Segmented Dataset towards Visual Detection of Marine Debris.*

6. *Wang, Tao & Cai, Yuanzheng & Liang, Lingyu & Ye, Dongyi. (2020). A Multi-Level Approach to Waste Object Segmentation.*

Thanks for the attention!