

Omdena Project: AI [Computer Vision] for Sorting and Segregating Waste Materials - Bengaluru Local Chapter

Task-1: Data Collection and Preprocessing

Resources for Data Sets:

1. <https://www.kaggle.com/datasets/asdasdasdasdas/garbage-classification>
 - 6 categories of waste segregation
2. [GitHub - AgaMiko/waste-datasets-review: List of image datasets with any kind of litter, garbage, waste and trash](#)
3. <https://github.com/garythung/trashnet>
 - (Used in Most of the Papers)
 - It has 6 classes: Glass, Paper, Cardboard, Plastic, Metal, Trash
 - It has 2527 Images
4. <https://github.com/cardstdani/WasteClassificationNeuralNetwork/tree/main/WasteImagesDataset>
5. <https://www.kaggle.com/datasets/techsash/waste-classification-data/code?datasetId=233210&sortBy=voteCount>
 - It has 2 classes: Recyclable and Organic Objects
 - It has 22500 Images.
6. <https://visualdata.io/discovery/dataset/abc7eb90870855be326d6f0abc54a91291135ff8>
 - <https://github.com/spotgarbage/spotgarbage-GINI/tree/master/spotgarbage/garbage-queried-images>
 - Garbage and Non-Garbage Labels
 - It contains 2561 images with 956 images containing garbage and the rest are non-garbage images.
7. <https://www.kaggle.com/datasets/dataclusterlabs/domestic-trash-garbage-dataset>
 - 9000+ original Trash/Garbage images
 - Labels Not checked (Not Labeled)
8. <http://tacodataset.org>
 - Open image dataset of waste in the wild.
 - Images are manually labeled and segmented according to a hierarchical taxonomy to train and evaluate object detection algorithms.
 - GH link for that: <https://github.com/pedropro/TACO>
9. <https://dockship.io/challenges/5f74153dfbaa7b0393ff19ea/garbage-classification-ai-challenge/overview>
 - 6 categories of waste segregation
10. <https://universe.roboflow.com/material-identification/garbage-classification-3>
 - It has 10464 images.
 - Classified into: BIODEGRADABLE, CARDBOARD, CLOTH, GLASS, METAL, PAPER, PLASTIC

11. <https://www.kaggle.com/mostafaabla/garbage-classification>

- This dataset has 15,150 images
- 12 different classes: paper, cardboard, biological, metal, plastic, green-glass, brown-glass, white-glass, clothes, shoes, batteries, and trash
- The clothes category and 22 % of the shoes' category were obtained from the Clothing dataset <https://www.kaggle.com/agrigorev/clothing-dataset-full>
- Around 29% of the other 9 classes combined was collected from the Garbage Classification dataset <https://www.kaggle.com/asdasdasdasdas/garbage-classification>
- All the rest of the images were obtained using Web Scrapping.

All the images collected using web scrapping are owned by their original photographers/owners. The images are only used to promote research and is not used to generate any profit. Any of the images in the data set can be deleted upon the request of the original photographer/owner.

12. <https://www.kaggle.com/datasets/techsash/waste-classification-data>

- Labeled Recyclable and Organic wastes that can be added to the decided 6 categories

Web Scrapping Resources:

1. <https://www.selenium.dev/documentation/>
2. <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
3. <https://pypi.org/project/requests/>

Readings for concepts/learning/knowledge/references:

1. <https://drive.google.com/drive/folders/1znvHurQdqDcmWxrkgUZjw0RkeZzveMZD>
2. https://www.biendata.xyz/competition/haihua_wastesorting_task1/evaluation/
3. https://www.biendata.net/competition/haihua_wastesorting_task2/evaluation/
4. Articles:
 - a. <https://medium.com/ramudroid/computer-vision-for-garbage-detection-136029142b3c>
 - b. <https://towardsdatascience.com/advanced-waste-classification-with-machine-learning-6445bff1304f> [Article]
5. Papers:
 - a. <https://asset-pdf.scinapse.io/prod/3045345459/3045345459.pdf>
 - b. http://cs230.stanford.edu/projects_spring_2020/reports/38847029.pdf
 - c. <https://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf>
 - d. https://www.researchgate.net/publication/325626219_RecycleNet_Intelligent_Waste_Sorting_Using_Deep_Neural_Networks
 - e. <https://core.ac.uk/download/pdf/148033278.pdf>

