Segregating Household Waste using Image-based Neural Networks

ANSHUL SANJAY DHARGAVE
M.Sc. in Electronic and Computer Technology
Dublin City University
Dublin, Ireland
anshul.dhargave2@mail.dcu.ie

***Abstract*—Owing to the drastic increase in solid waste and garbage generated by humans, waste management, segregation and recycling have become problems of scale. This warrants an automated solution, rather than relying on manual workforce, which will help in increasing life quality and give more control over the full waste generation chain from the source to the dump. This review paper dives into the idea of image-based waste segregation devices & approaches by looking at state-of-the-art studies, followed by an in-depth review of datasets, models and evaluation metrics to further investigate the idea and ascertain if this concept requires further exploration.**

***Index Terms*—Waste segregation, image-based deep learning, object detection, neural networks.**

I. Introduction

Household waste generated across the world is mostly disposed off in land-fills or entirely incinerated to avoid the cost of recycling [1]. One of the problems with the land-filling solution is that land is a critical resource and the amount of land convenient for transportation that can be allocated for waste disposal is very limited. Land-filling also affects the quality of life of the people and animals living around the land-fill area. Incineration, on the other hand, leads to increase in air pollution due to the high amount of plastics and other hazardous products involved. One of the ways to lower the strain on land-fills and to avoid combustion of hazardous materials is to recycle the waste items. For waste and discarded items to be recycled efficiently, they need to be segregated and classified into separate categories based on their raw materials, size, and use after recycling. The first step of this process would be to separate the biodegradable and the non-biodegradable waste items. This is feasible and even practised on smaller levels, but for city-level waste collection facilities, this becomes a problem of scale.

The best way to start the process of recycling is to start the waste segregation as early as possible. The ideal way to achieve this would be waste segregation at the source of waste generation. Such systems which allow households to segregate waste items as biodegradable, recyclable and non-recyclable, apart from glass and batteries, already exist in most developed countries. However, most developing countries lack the means of transporting the waste that was segregated at source either due to financial limitations or lack of interest [1], and so it is more prevalent to transport all the recyclable and non-recyclable waste items combined. This in turn requires the presence of a municipal or a city level waste segregation level before the land-filling stage.

Currently, most developing nations rely on manual waste separation and transportation for waste segregation [1]. This practice severely affects the quality of life of manual workers by increasing the risk of diseases and infections as the waste segregation sites are usually not sanitary and seldom involve stray animals & animal carcasses. The implementation of a scaled automatic waste segregation system that requires minimum human intervention would not only save the human lives but would also speed-up the waste segregation process.

Software systems that require an image input to complete their tasks are known as image-based systems, or more formally computer vision systems. Computer vision (CV) is also a stream of machine learning that focuses on development of automatic systems that 'trains' computers to interpret the input images, classify and recognise objects inside the input images, or even generate new images from scratch. This is done using neural networks (a combination of CV neurons), which are based on workings of a neuron in a human brain. The application of such systems include self-driving cars, traffic sign recognition systems for accessibility, face recognition systems, intruder detection systems, etc. Since image-based systems rely on image inputs, that can also be from a live camera, they can be applied on any camera-enabled 'smart'-device that requires analysis of the input image. Since the deployment of such a system can even be on-device, the cybersecurity and privacy concerns can also be easily addressed. This makes such technology a viable candidate for exploration of a solution for the waste segregation problem.

The final implementation of an image-based waste segregation system will differ nation-to-nation and city-to-city based on the availability of resources, feasibility of materials and implementation of ideas. This research aims to identify the feasibility of implementation and to then lay the foundations of an image-based waste segregation system using object detection neural networks by critically reviewing the existing literature.

II. Review and Analysis of Previous Work

Authors of [2] focus on just the Indian region for waste segregation. They propose an image-based system for waste segregation that is based on a custom Convolutional Neural

Network (CNN) that uses the *Trashnet* dataset for training and testing the model. The dataset consists of four usable categories of waste items, namely paper, metal, plastic and cardboard, with a total of 2077 images. The supervised learning CNN classifier used in this paper manages to get a 76% accuracy on the labelled dataset. The authors of [2] note that the accuracy is less and claim that the limited number of images is the cause. They also claim that with improved accuracy their model can be used for home automation and practical applications. They, however, fail to state any false positives or false negatives that their model may have classified. They also fail to mention any time limitations that may occur during processing when their model is used in practical applications. [2] shows that the number of images and the choice of dataset will play an important role in direction and implementation of such a waste segregation system.

[3] is a recent study that proposes a better approach compared to [2] by employing the Xception paradigm-based neural networks for their waste classification. [3] use the *Garbage Collection dataset* extracted from the famous *Imagenet* dataset which has 12 different waste item classes for more controlled waste classification. In [3], the authors detail that the implementation part of the study was done on *Google*'s *Colab* framework using *Python*'s *Keras* library and *TensorFlow* backend. The authors demonstrate a testing accuracy of 96% using cross-validation technique and splitting the training dataset into 80:20 ratio, where 80 is the training set ratio and 20 is the validation set ratio. They also note that further optimising their proposed model can result in lower latencies while classifying trash items. One way in which this study can be improved is by further constraining the choice of training image dataset. The dataset used here has an imbalanced class that could affect the segregation process once the model is deployed 'in-the-wild'.

[4] explore the usability of CV-based models for waste segregation by comparing various pre-trained neural networks on the *TrashNet* dataset. The *TrashNet* dataset contains six waste item classes with 2527 usable images, each of resolution 512x384. They compared five pre-trained neural networks, namely VGG-16, VGG-19, Inception, ResNet, and Inception-ResNet. The highest accuracy they achieved was 88.66% using the ResNet model. The authors of [4] note that they have achieved the highest accuracy on the same dataset compared to the state-of-the-art, which also includes [5]. Although the best in terms of accuracy, [4] agree that the best way to improve the current state of models would be to use a realistic synthetic image dataset mixed with real images from the 'wild'. The real-images test factor will play a huge role in terms of deployment of the proposed model.

[6], like [2], focus on the Indian context for their waste segregation study. They authors of [6] note that the Indian government, to ease the waste segregation at the household level, have introduced wet, dry and mixed waste categories. [6] focuses on this relatively simplified tri-class system to develop an *Android* app on CV-based model to enable citizens to segregate waste items at their household level. The dataset used in this study is a combination of two datasets. For the non-biodegradable class, the *Trashnet* dataset is used, whereas for the biodegredable class, the authors of [6] have used their own images. This combination results in a total of 2700 usable images for the training of model. The CNN model used in this study is InceptionV3 network that was trained on the *Colab* framework by *Google* using the *TensorFlow* library. The model achieved an accuracy of 88% on biodegradable item images with 1.3 seconds execution time and 84% accuracy on non-biodegradable item images with 1.56 seconds of execution time, which makes it usable in real-time applications.

[7] also uses a deep learning YOLOv3 network to help waste segregation using computer vision. The dataset used in this study was created by them using an iPhone. The total number of usable images after the pre-processing stage were 6437 across six classes of different waste items, namely cardboard, glass, metal, paper and plastic. [7] compares the accuracies and prediction times of the YOLOv3 and YOLOv3-tiny neural networks. It finds that the YOLOv3 has higher accuracies than YOLOv3-tiny across all classes but also takes almost four times the execution times of YOLOv3-tiny. Apart from lower accuracies, YOLOv3-tiny also suffers from high false detections and a few non-detection scenarios. This is mainly because YOLOv3-tiny has a simplified architecture as compared to YOLOv3 which helps in deployment on edge-devices. This study, although shows interesting comparison of accuracy vs execution times, does not mention the optimisation that went into improving the YOLOv3-tiny network's performance.

[8] proposes a hardware based solution coupled with the deep learning methods to improve waste segregation at the source. [8] describes the instrumentation required to build such a system along with the CV-based models that run the system. The authors of [8] have made sure to make the device as affordable as possible, with the total cost of all materials coming close to 4000 Indian Rupees (or approximately US$50). The final dataset used in the paper is combined from three separate sources. The first dataset source is the *Trashnet* dataset for the five classes of glass, cardboard, paper, plastic, metal and others. The second dataset source is the *Waste Classification* data which provided two classes, organic and recyclable items. The third dataset source is the *Drinking Waste* classification dataset containing 4 classes of glass, aluminium cans, PET and HDPE. These datasets combine to give 9516 images with biodegradable and non-biodegradable as 51.76% and 48.24% respectively. The models used to implement this system are AlexNet (97.95% accuracy), ResNet (97.21% accuracy), InceptionV3 (96.23% accuracy), and VGG-16 (87.52% accuracy). Although the accuracies quoted in this study are higher than the state-of-the-art, the authors of [8] do not mention the false detections and the execution times for all these models on the final implementation of the device. Furthermore, the image dataset generated had a completely white background, which will rarely be the case in real waste images, so a mixed approach with the training images would have benefited more in terms of practical applicability.

[9] propose a waste classification system using the ResNet network-based pretrained CNN model and the Support Vector Machine (SVM) machine learning model. The ResNet CNN is used as an extractor and the SVM system is used as a

classifier. The dataset used for this study is the *Trashnet* dataset with 1989 images, each with resolution of 512x384, belonging to four classes - glass, paper, plastic and metal. The final implementation of the model works such taht the ResNet model is trained on the images and the weights learned by the model on the *Trashnet* dataset are then transferred to the SVM model, which classifies the input image into the four categories. The choice of SVM as the classifier is justified in [9] by mentioning that it is one of the best classifiers in machine learning. The final model gets an accuracy of 87%. Results in [9] are just based on the metric of accuracy. There is no mention of any false detections, no detections or the execution time. The authors of [9] agree that the result can be improved by using a larger dataset with more high-quality and real images.

[5], like [6], aims for the waste segregation at source and showcases a mobile app based on CV applications. The proposed app depends on citizens to track and report on their surroundings by uploading an image of garbage items using their mobile's camera, which also sends the geographical information from their mobile phone. The dataset here is actually introduced by the authors of [5], called the *Garbage in Images* (GINI) dataset, which was created by scraping the web using Bing Image Search API resulting in 2561 images. The dataset used in [5] is the only dataset that also contains images for heaps of garbage and trash items so as to train the models to recognise a large quantity of waste accumulated in a location. The images in this dataset were all manually annotated into separate categories. [5] compares four CNN approaches: (1) HOG + Gabor + Colour approach that results in about 80% accuracy with approximately 25 seconds execution time, (2) Sliding Window CNN model that results in 87.21% with approx. 50 seconds execution time, (3) GarbNet CNN model with Local Response Normalisation (LRN) approach which results in 87.70% accuracy and 4.11 seconds execution time, and (4) GarbNet CNN model without LRN approach which results in 87.69% accuracy and 1.50 seconds execution time. Although the accuracy in [5] is not the highest, the execution is the most thorough with state-of-the-art models, datasets, and equipment. The specificities of approx. 80% for the first approach and approx. 90% for the rest of the approaches also suggest a low false detection rate. All the approaches also have 81-83% sensitivity to tackle the no-detection scenarios. The final mobile app developed in [5] was also tested on several *Android* phones popular at the time of publication, giving a sub-six second execution time and using 83% of CPU cores with 67.3MB of memory usage, which is usable in real-life scenarios, especially with the dramatic increase in the processing power of *Android* phones since the publication of this paper. The authors of [5] quote a final accuracy of 87.69% using the last approach using Garbnet CNN without LRN.

III. Relation of Prior Work to the Project Problem

The research reviewed in this work have been summarised in the I. From the review, the key points that affect the accuracy of the models, and also the execution time, are the choice of dataset, quality and quantity of images in the dataset, robustness of the machine learning model, architecture of the model, and finally the manner of execution of the proposed systems.

TABLE I
A SUMMARY OF ALL PAPERS REVIEWED FOR WASTE SEGREGATION USING IMAGE-BASED TECHNIQUES

Paper	Dataset	Methodology	Accuracy (%)
[2]	Trashnet	Custom CNN	76
[3]	Imagenet	XceptionNet	96
[4]	Trashnet	ResNet	88.66
[6]	Combination	InceptionV3	86
[7]	Custom	YOLOv3, YOLOv3-tiny	94.99, 51.95
[8]	Combination	AlexNet, ResNet	97.95, 97.21
[9]	Trashnet	ResNet + SVM	87
[5]	GINI	Garbnet	87.69

In the review, it was noted that [5] had the best quality dataset and the most robust testing of models. Out of all the papers reviewed, [5] is the only study that includes the provision to classify accumulated waste along with separate waste items. This allows for a better real-life application scenario where individual waste items are seldom the case. [7] showcased that the choice of models depending on their architecture, whether the model has many layers or whether it has a simplified architecture, yield very different results in terms of accuracy and execution times. [3] and [6] showcase that while state-of-the-art equipment used in [5] helps in more all-rounded testing, *Google*'s free cloud-based *Colab* platform can be used for individual research to yield respectable results. [2], [4], [6] suggest that the use of the *Trashnet* dataset using the various pre-trained CNNs result in approximately similar accuracies. [5], [8] confirm that using images from different sources and combining them yields the best results as the context is more comprehensive, and hence the application is closer to real-life scenarios. From all the papers, it is also clear that execution time and false detection (along with no detection) play an important role as evaluation metrics, apart from the usual classification accuracy metric, and help paint a wider picture.

IV. Conclusion

After careful and critical review of several state-of-the-art studies, it is concluded that the concept of waste segregation using image-based CV techniques have not been fully explored, and hence this idea can be investigated further. The choice of dataset(s) and model(s) will be constrained in further research. So far, the *Trashnet* and *GINI* datasets have shown to be quite comprehensive. For model choices, interim research suggests pitting pre-trained CNNs like Inception and ResNet against custom CNNs like Garbnet and YOLOv3. During analysis of evaluation metrics, it was also noted that accuracies do not always show the complete truth, several other metrics, such as execution time and false/no detections, should also be discussed.

REFERENCES

[1] D. C. Wilson, L. Rodic, A. Scheinberg, C. A. Velis, and G. Alabaster, "Comparative analysis of solid waste management in 20 cities," *Waste Management & Research*, vol. 30, no. 3, pp. 237–254, 2012, pMID: 22407700. [Online]. Available: https://doi.org/10.1177/0734242X12437569

[2] S. R., R. P., V. S., K. R., and G. M., "Deep learning based smart garbage classifier for effective waste management," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, 2020, pp. 1086–1089.

[3] T. N. N. H. ALilly Raamesh, S. Kalarani, "A methodical and intuitive image classifier for trash categorization based on deep learning," *International Journal of Modern Agriculture*, vol. 10, no. 2, pp. 4638 – 4652, Aug. 2021.

[4] V. Ruiz, Á. Sánchez, J. F. Vélez, and B. Raducanu, "Automatic image-based waste classification," in *From Bioinspired Systems and Biomedical Applications to Machine Learning*, J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, J. Toledo Moreo, and H. Adeli, Eds. Cham: Springer International Publishing, 2019, pp. 422–431.

[5] G. Mittal, K. B. Yagnik, M. Garg, and N. C. Krishnan, "Spotgarbage: Smartphone app to detect garbage using deep learning," p. 940–945, 2016. [Online]. Available: https://doi.org/10.1145/2971648.2971731

[6] F. Shaikh, N. Kazi, F. Khan, and Z. Thakur, "Waste profiling and analysis using machine learning," pp. 488–492, 2020.

[7] S. Kumar, D. Yadav, H. Gupta, O. P. Verma, I. A. Ansari, and C. W. Ahn, "A novel yolov3 algorithm-based deep learning approach for waste segregation: Towards smart waste management," *Electronics*, vol. 10, no. 1, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/1/14

[8] T. Gupta, R. Joshi, D. Mukhopadhyay, K. Sachdeva, N. Jain, D. Virmani, and L. Garcia-Hernandez, "A deep learning approach based hardware solution to categorise garbage in environment," *Complex Intelligent Systems*, vol. 8, no. 2, p. 1129–1152, Apr 2022. [Online]. Available: https://doi.org/10.1007/s40747-021-00529-0

[9] O. Adedeji and Z. Wang, "Intelligent waste classification system using deep learning convolutional neural network," *Procedia Manufacturing*, vol. 35, p. 607–612, Jan 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2351978919307231