Classification of Organic and Solid Waste Using Deep Convolutional Neural Networks

Conference Paper · September 2021

DOI: 10.1109/R10-HTC53172.2021.9641560

4 authors:

Rushnan Faria
Chittagong University of Engineering & Technology

Fahmida Ahmed
Chittagong University of Engineering & Technology

Annesha Das
Chittagong University of Engineering & Technology

Ashim Dey
Chittagong University of Engineering & Technology

Classification of Organic and Solid Waste Using Deep Convolutional Neural Networks

Rushnan Faria*, Fahmida Ahmed†, Annesha Das‡ and Ashim Dey§
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chittagong-4349, Bangladesh
*u1604110@student.cuet.ac.bd, †u1604107@student.cuet.ac.bd, ‡annesha@cuet.ac.bd, §ashim@cuet.ac.bd

***Abstract*—The total amount of waste is increasing all around the world day-by-day especially in urban areas. The increasing amount of unprocessed waste is very dangerous to mankind as it creates severe pollution in the environment. Most of this wastage is recyclable. For recycling, the waste needs to be separated at first, as different types of waste require different recycling techniques. But unfortunately, categorizing waste manually is very costly and time-consuming. So, in this work, a method is proposed to automatically classify waste into four categories. For this, a dataset named *OrgalidWaste* is prepared by collecting images from four other waste datasets. The prepared dataset contains around 5600 images with four classes including one organic waste class and three solid waste classes (glass, metal, and plastic). On this dataset, several CNN architectures including 3-layer CNN, VGG16, VGG19, Inception-V3, and ResNet50 have been implemented for training. Among them, VGG16 outperforms other models with 88.42% accuracy. It is believed that this work will be greatly beneficial in the waste management sector.**

***Keywords*—Recycling, Waste classification, Organic waste, Solid waste, CNN, VGG16**

I. Introduction

A large amount of waste is produced by humans all over the world every day. This waste includes mainly five types which are solid, liquid, organic, recyclable, and hazardous. 2.01 billion tons of municipal solid waste is produced by people yearly and at least 33% of this waste is not handled according to environmental safety regulations. Moreover, each person generates 0.74 kilograms of waste on average per day which varies greatly from 0.11 to 4.54 kilograms worldwide. Although only 16% of the world's population belongs to high-income countries, they produce around 34% or 683 million tons of the total waste around the world [1]. The amount of waste production is increasing day by day. In September 2018, the World Bank predicted that if no urgent action is taken for waste management soon, global waste production will be risen by 70% within 2050 [2]. Furthermore, this large amount of wastage is a great threat to our environment if not properly managed. Recycling can be a good solution to this problem. Different types of wastage need different types of recycling processes. But most of the time, all types of garbage are combined in a single bin. Governments of many countries are taking the necessary steps to reduce this problem by providing separate bins for different types of garbage. But the problem is that people do not follow these rules strictly and also, human error is inevitable.

Without classifying the garbage, it becomes immensely difficult to recycle them. So, classification is necessary before recycling them according to their required process. But doing this manually is very time-consuming, difficult and also it needs a large amount of manpower and money. To overcome this problem, an automatic categorizer can be a reliable and cost-effective solution. Recent advancements in deep learning have made a breakthrough by improving the accuracy in the area of image classification. The deep learning model passes queries through various hierarchies of concepts to find answers, which is similar to the human brain. In recent years, many researcher have proposed variety of solutions utilizing deep learning models (Resnet18 [3], Resnet50 with Support Vector Machine (SVM) [4], Inception-V3 [5]) and hardware implementation. But most of these works have classified waste into biodegradable and non-biodegradable ones and some other works have only classified solid materials. In the case of biodegradable, recycling processes are the same for all. But in the case of non-biodegradable which includes glass, metal, plastic, etc., different recycling processes may be needed.

So, considering all these, the main aim of this work is to classify garbage into four categories. Among these four categories, one is organic and the rest are solid materials including glass, metal, and plastic. To classify them, a large dataset is prepared by combining four datasets and then implementation of several deep learning models for training is done. There are some waste classification datasets available that include either solid material classes (glass, plastic, metal, etc.) or biodegradable and non-biodegradable classes. To the best of the authors' knowledge, there is no single dataset available that consists of organic, plastic, glass, and metal classes for waste classification. The main objectives of this work are:

- To prepare a large dataset containing one organic waste class and three solid waste classes (glass, metal, and plastic)
- To implement various deep learning models for training them using our dataset.
- To compare and analyze the performance of these models based on accuracy.

The rest of this paper is outlined as follows. Section II is presenting the related works. The proposed methodology is described in Section III. Results are presented in Section IV. And the paper is concluded in Section V.

II. Literature Review

Waste classification has become one of the most popular research topics because of the increasing amount of unprocessed waste around the world. For classifying waste, many research works have been done using various types of datasets and machine learning techniques.

In [3], TrashNet dataset was used and different Convolutional Neural Network (CNN) models like VGG16, VGG19, Inception, ResNet18, and Inception-ResNet were implemented. In their experiment, they achieved the highest accuracy of 88.66% for ResNet18.

In [4], an intelligent waste classification system was proposed where classification was done into four classes including glass, paper, metal, and plastic. To train and test the model, they used a trash image dataset. The system was developed by using the 50 layers residual network CNN model as the extractor and SVM as the classifier. This proposed system achieved 87% accuracy.

In [5], authors worked with office garbage. A dataset was created by collecting pictures from the internet and taking images with a Raspberry Pi wide-angle camera. There were six classes in their dataset including battery, milk-box, can, paper, paper cup, and bottle. Higher accuracy of 95.33% was achieved using Inception-V3 CNN model.

In [6], authors prepared TrashNet dataset for waste classification. This dataset has six classes including trash, paper, metal, plastic, glass, and cardboard. In this work, a comparison between SVM and CNN has been shown where SVM achieved 63% and CNN achieved 22% accuracy.

In [7], another research was done on bulky waste classification. A dataset was created which includes 95 classes of the bulky waste and each of these classes contains 500 images. In this work, they implemented a fine-tuned VGG19 model and three hybrid models to handle imbalance data. In this work, the proposed hybrid model obtained an accuracy of 86.19%.

In [8], authors focused on separating biodegradable and non-biodegradable waste. Here, the training model was made using inception-V3 architecture. The model achieved an accuracy of 83.30%.

In [9], Lulea University of Technology developed a project to recycle metal garbage by a mechanical identifier. The system identifies the chemical contents and the actual separation by using chemical and mechanical methodologies.

In [10], an image-based classifier was built using a dataset of materials collected from Flickr, the well-known image sharing service. Their system can identify the fabrics of the materials category from a single image of a surface using a Bayesian Classifier. It achieved a 44.6% recognition rate.

In [11], authors developed a system that analyzed different types of real-world waste material captured by Nearest Infrared Ray (NIR) and RGB cameras. This experiment showed that to correctly classify images, the NIR information is adequate. But, in some cases, problems could arise to classify objects on the conveyor belt. For solving this problem, they exploited an algorithm that integrates two different sensors' information, resulting in right classification in the most tricky cases.

In [12], the authors presented a deep learning model named AquaVision which detects and classifies waste floating in the water. The different floating waste on the seashores and in the oceans are detected and classified with mean Average Precision (mAP) of 0.8148 by this model. A dataset was created named AquaTrash merging two existing datasets named TrashNet and TACO. They proposed a RetinaNet which used ResNet50 and Feature Pyramid Network (FPN) as the backbone.

In this work, a dataset is prepared named *OrgalidWaste* containing about 5600 images with four classes of waste including one organic waste class and three solid waste classes (glass, metal, plastic). The dataset was trained using 3-layer CNN, VGG16, VGG19, Inception-V3, and ResNet50.

III. Methodology

This work aims to categorize waste into four waste classes (organic, glass, metal, and plastic) from images of recyclable and biodegradable waste. Fig. 1 shows the overall workflow of this research.

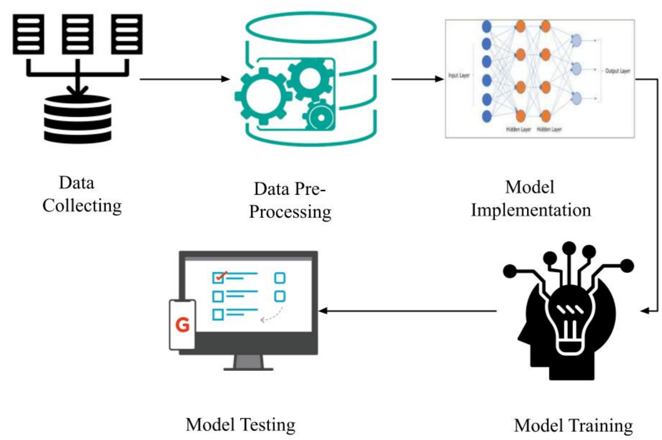

Fig. 1. Overview of work flow.

To classify waste, four datasets are collected then a combined dataset is prepared using them. Then, preprocessing is done using augmentation. After that, several CNN models are trained with the generated dataset and compared based on their accuracy. The overall methodology is described in the following sub-sections:

A. Dataset Preparation
B. Dataset Preprocessing
C. Model Implementation
D. Training and Testing

Each of these sub-sections is described below-

A. Dataset Preparation

In this step, four datasets are used to create a new dataset. Images of different classes were collected from these datasets and merged into *OrgalidWaste*. All these datasets are described as follows-

1) Drinking waste classification dataset [13]: This dataset was manually labeled and collected by Arkadiy Serezhkin. The format of images in this dataset is JPG. This is publicly available on the Kaggle website. The size of this dataset is 1GB including raw images and YOLO images folder. In raw image folder, it has 4828 images. Besides, the YOLO images folder consists of raw image folder binding with a text file for the YOLO framework. Images from the raw image folder are used that consists of four classes named Glass (1232 images), AluCan (1060 images), PET (1028 images), and HDPEM (1508 images). Fig. 2 represents that the dataset is balanced. The class named Glass consists of pictures of glass bottles, AluCan consists of aluminum cans, PET consists of plastic bottles and HDPEM consists of plastic milk bottles.

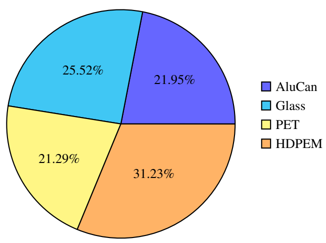

Fig. 2. Distribution of drinking waste classification dataset.

2) Waste classification dataset [14]: This dataset was created by Sashaank Sekar which is available on the Kaggle website. The total size of this dataset is 427MB and all the images are in JPG format. In this dataset, the images are divided into train and test folders. In each folder, there are two subfolders one is organic and another is recyclable. In total, they have 22500 images among which organic class has 14001 images and recyclable class has 11111 images. Fig. 3 represents that the dataset is almost balanced. The organic class contains images of various vegetables, fruits, meats, eggs, bakery items, etc, and the recyclable class contains images of glass, plastic, metal, cardboard, paper, books, cooking pot, etc.

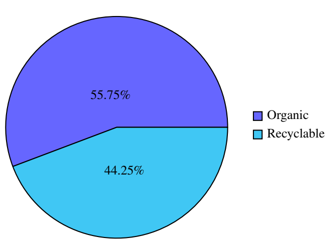

Fig. 3. Distribution of waste classification dataset.

3) TrashNet [6]: Gary Thung and Mindy Yang created this dataset which is publicly available on different websites (Kaggle, Github). The size of the dataset is 82MB and all the images are in JPG format. This dataset contains six categories. They are glass, metal, plastic, cardboard, paper, and trash. This dataset consists of images of glass, bottles, boxes, cans, lids, torn newspapers, magazines, cardboard images including torn cardboard, packets, etc. TABLE I presents detailed information about this dataset.

TABLE I
DETAIL OF TRASHNET DATASET

Classes	***Details***	***Total Data***
Cardboard	Card board boxes, Torn cardboard	393
Glass	Glass, Bottles, Jar, Broken bottles	491
Metal	Can, Smashed Can, Jar lid	400
Paper	Newspaper, Page of Magazines, Envelope, Paper	584
Plastic	Bottles, Box, Milk Bottle	482
Trash	Different types of packets like chocolate packets, chips packets, Sauce packets	127

4) Trash dataset [15]: This dataset was created by Dimon Is Ochecolo. The size of this dataset is 92MB and is publicly available on the Kaggle website. The total number of images in the Trash dataset is 3907 among which 136 images are in PNG format and 3771 images are in JPG format. This dataset consists of four classes including glass, paper, metal, and plastics. Glass class has pictures of glass, bottles, broken glass, etc. The plastic class contains pictures of bottles, plastic bags, plastic bottle caps, etc. The metal class includes the images of metal cans, lids, etc., and the paper class includes the images of torn paper, juice packets, cardboard, etc. The detailed information of this dataset is given in TABLE II.

TABLE II
DETAIL OF TRASH DATASET

Classes	***Details***	***Total Data***
Glass	Glass, Bottles, Vase, Jug	914
Metal	Can, Smashed can, bottle's cap	801
Paper	Card Board, Packet of Juice, Cup	1187
Plastic	Bottles, Plastic bag, Packets, Bottles Cap	1005

5) OrgalidWaste *dataset:* This dataset is named as '*OrgalidWaste*' which is derived from the words organic and solid. In this dataset, there are four classes including one organic class, and three solid classes (glass, metal, and plastic). The dataset is split into train, validation, and test. The total size of this dataset is 93MB with around 5600 images. This dataset has 20% data for validation and 10% for the test section. TABLE III shows the detailed information of the dataset. All the images are in JPG format. Fig. 4 shows some of the sample images from each class of this dataset. From the drinking

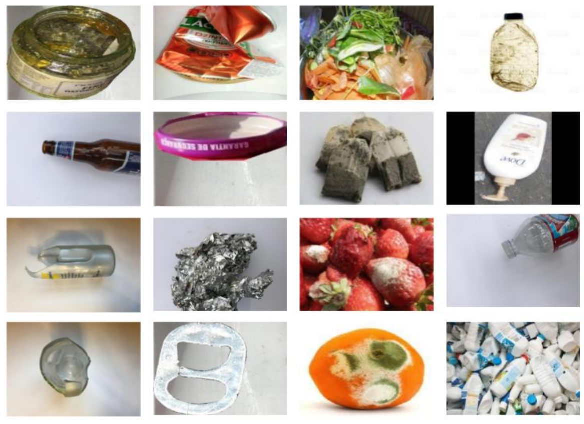

Fig. 4. Sample images from *OrgalidWaste* dataset.

TABLE III
DETAIL OF *OrgalidWaste* DATASET

Classes	*Details*	*Data*
Glass	Broken Bottles, Jar, Broken Glass, Drinking Glass	1415
Metal	Smashed Can, Lid, Aluminium foil, Coin, Key, Cooking pot, Spoon and Fork	1413
Organic	Different Rotten Vegetables and Fruits, Eggshell, Tea bag, Meat, Peel	1435
Plastic	Milk Jar, Smashed Bottle, Bottle's cap, Spoon, Box, Water Pot, Packets	1420

waste classification dataset, several images are taken for the glass, metal, and plastic classes of this dataset. Organic images and also some images of glass, metal, and plastic are collected from the waste classification dataset. Some images of plastic, glass, and metal are also taken from both TrashNet and Trash datasets. The prepared dataset is balanced in terms of the number of images in each class which is around 1400.

B. Data Preprocessing

Data augmentation is carried out in the *OrgalidWaste* dataset which is beneficial to enhance performance and end results of machine learning models by building different and new examples to train datasets. For this, the function named '*ImageDataGenerator*' has been used from Keras.preprocessing package. This function has several parameters among which the used arguments are 20-degree rotation, horizontal flip, vertical flip, 0.1 width shift, and 0.1 height shift. The images are resized and 224px for image height and width are taken.

C. Model Implementation

To do the classification, five architectures of CNN have been implemented using Google Colab environment which provides Nvidia Tesla K-80 GPU totally free. Among them, four architectures named VGG16, VGG19, ResNet50, and Inception-V3 are utilized using transfer learning. Transfer learning is a 'design methodology' within machine learning. Transfer learning [16] extracts helpful information from data in a relevant field and solves cross-domain learning problems. Transfer learning can be a great solution when the machine learning model can be over-fitted. In transfer learning, the learning process will start from patterns that have learned to solve a different task rather than starting with a blank sheet (often randomly initialized). All five architectures are deep learning models. The deep learning model has many layers and sends the input through different layers. At the time of the training, each of these layers learns specific features from images, and the network can pull out these features at any time.

Details about the five architectures are described below:

1) 3-layer CNN : CNN has some convolutional layers in which a convolution function connects the neurons. The convolutional layer, the ReLU correction layer, the pooling layer, and the fully connected layer are the four classes of the CNN layer. Three convolutional layers are used with three max-pooling layers where the pool size was 2*2. ReLU activation functions are also used in these CNN layers. And then, the fully connected layer has a flatten, dense, and dropout layer. The output layer has the softmax activation function and four nodes.

2) VGG16: VGG16 [16] has a block of 13 convolution layers and 3 fully connected layers. It does not have a large number of hyper-parameter. It focuses on 3*3 convolution layer filters and 2*2 max-pooling filters. Softmax follows the fully connected layers for output.

3) VGG19: VGG19 [16] is a variation of VGG16. There is only one distinction between these two architectures and that is the last three convolutional blocks consist of 4 convolutional layers in VGG19 whereas VGG16 consists of 3 convolutional layers.

4) ResNet50: Residual Network (ResNet) comes in the scenario for activating hundreds or thousands of convolutional layers. The vanishing gradient problem can be solved by this architecture using identity shortcut connections. ResNet50 is a variant of the ResNet model. It consists of 48 convolutional layers with one average pool and one max pool layer. It has 3800 million floating points operations.

5) Inception-V3: Inception-V3 consists of 48 deep layers. This model has 11 inception modules. Every module contains pooling layers and convolution filters with ReLU as the activation function. This model contains label smoothing, factorized 7*7 convolutions, and an auxiliary classifier.

For implementing the transfer learning models (VGG16, VGG19, ResNet50, and Inception-V3), the same steps are followed. At first, these models were trained using the ImageNet dataset. ImageNet [17] is an image database that is arranged following the WordNet hierarchy. Every node of the hierarchy is presented with hundreds and thousands of pictures. These models' output labels are kept false and the weights are used

TABLE IV
ACCURACY COMPARISON

Architecture	*Accuracy (%)*	*Validation Accuracy (%)*	*Epoch*	*Execution Time (Minutes)*
3-layer CNN	80.31	80.71	50	92
VGG16	88.42	87.32	20	52
VGG19	86.38	85.98	25	54
ResNet50	50.28	48.04	20	46
Inception-V3	69.94	69.11	86	105

in a model that has a flatten layer, a dense layer containing 1024 nodes, and a dropout layer with 0.2 values. Finally, the model has the output layer containing 4 nodes and the softmax function as the activation function.

D. Training and Testing

For training our models, the Adam optimizer was applied. Categorical cross-entropy was applied as a loss function. The batch size is 32. To avoid overfitting, different epochs for different models were utilized. The epochs are set to 50 for three-layer CNN, 20 for VGG16, 25 for VGG19, 86 for Inception-V3, and 20 for ResNet50. For validating these models, the validation folder was used which contains about 1120 images. For testing these models, the test folder of the *OrgalidWaste* dataset is used that contains 580 images in total. An image is randomly taken from the test folder and then the prediction label and actual label for this image are compared. The '*randint*' function was used to select an image from that folder. The result is discussed in the following section.

IV. RESULT

In this section, the performance of five CNN models trained on the *OrgalidWaste* dataset has been presented. The accuracy and execution time of the different CNN architectures that are implemented is shown in TABLE IV. The accuracy and validation accuracy is obtained from the training and validation folder of the *OrgalidWaste* dataset. As it is known that the less the difference between accuracy and validation accuracy is, the more well learned the model is. Different epochs for different models are used to achieve lesser variance between accuracy and validation accuracy. Otherwise, the model becomes under-fitted or overfitted. Besides, data augmentation is used to avoid overfitting as well as underfitting problems, as well as to improve the model's performance.

From TABLE IV, it can be observed that VGG16 gained the highest accuracy among all the models, and the accuracy is 88.42%. The lower depth of the used dataset may account for VGG16's better performance. VGG19 also performs better compared to other implemented models with an accuracy of 86.38%. Fig. 5 shows the learning curve of the VGG16 model. At first, the epoch number is increased to avoid the underfitting problem. Furthermore, the overfit problem is taken into account when determining the epoch number. The epoch for this model is 20 because after 20 epoch accuracy and validation accuracy started to increase again which is not acceptable. Some zig-zag in the validation curve is seen. In general, this may occur because of the presence of noisy data. As the data is merged from different datasets, this could be a reason for noisy data. Apart from this, the glass and plastic class could be confusing with only RGB images. Fig. 6 shows the predicted label using VGG16 for sample images taken from the test dataset.

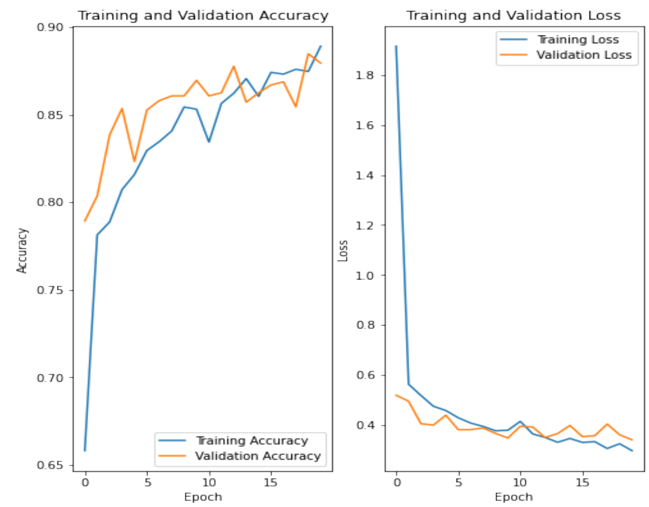

Fig. 5. Learning curve for VGG16.

TABLE V shows the confusion matrix values that describe the performance of the VGG16 model on the test folder of the *OrgalidWaste* dataset. The test folder consists of a total of 580 images. The diagonal values of the table show the rate of prediction for which the predicted class is as same as the true class. The higher diagonal values are the indication of a more correct prediction. It can be noticed from the diagonal of the matrix that every category is classified with a high prediction rate. The normalized form of the confusion matrix is used that visualizes more accurately the mispredicted classes. Metal and organic classes have a higher prediction rate compared to glass and plastic classes. And misprediction rates are very low for all classes.

TABLE V
CONFUSION MATRIX FOR VGG16

Class Labels	**Prediction**			
	Glass	***Metal***	***Organic***	***Plastic***
Glass	**0.73**	0.18	0	0.08
Metal	0.05	**0.93**	0.01	0.01
Organic	0.01	0.04	**0.95**	0.01
Plastic	0.13	0.09	0	**0.78**

V. CONCLUSION

The increasing amount of unprocessed waste around the world imposes great danger to the environment and nature. For the waste recycling process to be more effective, categorizing waste is very important. In this paper, The *OrgalidWaste* dataset is prepared that consists of around 5600 images using

Fig. 6. Predicted classes for sample images from the test dataset.

four publicly available waste datasets. The dataset has four waste classes which are organic, glass, metal, and plastic. Several CNN architectures including 3-layer CNN, VGG16, VGG19, Inception-V3, and ResNet50 have been implemented for training. Among them, VGG16 achieved the best result with 88.42% accuracy. In the future, it is aimed to enrich the *OrgalidWaste* dataset with more real-world waste images. This work can be beneficial to reduce the workload of waste management companies. Also, an automatic waste collector robot can be developed using this model embedded with different hardware.

References

[1] Worldbank, "Trends in solid waste management," https://datatopics.worldbank.org/what-a-waste/trends_in_solid_waste_management.html#:~:text=The%20world%20generates%202.01%20billion,from%200.11%20to%204.54%20kilograms, accessed July 5, 2021.

[2] Sensoneo, "Waste index 2019," https://sensoneo.com/sensoneo-global-waste-index-2019/, accessed July 5, 2021.

[3] V. Ruiz, Á. Sánchez, J. F. Vélez, and B. Raducanu, "Automatic image-based waste classification," in *International Work-Conference on the Interplay Between Natural and Artificial Computation*. Springer, 2019, pp. 422–431.

[4] O. Adedeji and Z. Wang, "Intelligent waste classification system using deep learning convolutional neural network," *Procedia Manufacturing*, vol. 35, pp. 607–612, 2019.

[5] J.-w. Feng and X.-y. Tang, "Office garbage intelligent classification based on inception-v3 transfer learning model," in *Journal of Physics: Conference Series*, vol. 1487. IOP Publishing, 2020, p. 012008.

[6] M. Yang and G. Thung, "Classification of trash for recyclability status," *CS229 Project Report*, vol. 2016, 2016.

[7] H. Wang, Y. Li, L. M. Dang, J. Ko, D. Han, and H. Moon, "Smartphone-based bulky waste classification using convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 39, pp. 29 411–29 431, 2020.

[8] F. Shaikh, N. Kazi, F. Khan, and Z. Thakur, "Waste profiling and analysis using machine learning," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, 2020, pp. 488–492.

[9] S. Zhang and E. Forssberg, "Intelligent liberation and classification of electronic scrap," *Powder technology*, vol. 105, no. 1-3, pp. 295–301, 1999.

[10] C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz, "Exploring features in a bayesian framework for material recognition," in *2010 IEEE computer society conference on computer vision and pattern recognition*. IEEE, 2010, pp. 239–246.

[11] E. Strollo, G. Sansonetti, M. C. Mayer, C. Limongelli, and A. Micarelli, "An AI-Based approach to automatic waste sorting," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 662–669.

[12] H. Panwar, P. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, S. Sharma, and I. H. Sarker, "Aquavision: Automating the detection of waste in water bodies using deep transfer learning," *Case Studies in Chemical and Environmental Engineering*, p. 100026, 2020.

[13] A. Serezhkin, "Drinking waste classification," https://www.kaggle.com/arkadiyhacks/drinking-waste-classification, accessed July 5, 2021.

[14] S. Sekar, "Waste classification data," https://www.kaggle.com/techsash/waste-classification-data, accessed July 5, 2021.

[15] D. I. OCHECOLO, "trash_dataset," https://www.kaggle.com/dimonisochecolo/trash-dataset, accessed July 5, 2021.

[16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.

[17] "Image net," http://www.image-net.org/, accessed July 5, 2021.