Eco-Bin: Automated Trash Sorting for Environmental Sustainability

Mehul Goel, Anirudh Mani, Arul Rhik Mazumder and Steven Yang

July 30, 2024

---

1 Introduction

Waste management is an essential but often overlooked sector. Large accumulations of urban waste not only threaten the environment by releasing greenhouse gases and toxic pollutants but also pose a risk to humans by attracting rats, mosquitoes, and other disease vectors. Despite the current management systems, waste control challenges will only continue to worsen due to the increasing population and lack of attention to the problem. At the moment, current urban centers produce about 2 billion tons of waste, but by 2050, this number is projected to rise to 3.4 billion [Lahcen et al., 2022]. About 70% of this waste will be mismanaged and end up in landfills to pollute the earth [Poudel and Poudyal, 2023].

One of the greatest solutions to the waste accumulation problem is recycling. Recycling reduces waste accumulation by reusing specific waste materials and is a key process in developing a sustainable circular society that minimizes waste and reduces excess resource extraction and consumption [Nnamoko et al., 2022]. Unfortunately, current recycling procedures are primitive and inefficient, often carried out manually by employed sorters. This process is both inefficient and dangerous for the waste sorters, who risk exposure to diseases or toxic contaminants [Adedeji and Wang, 2019]. Eco-Bin is a novel machine learning support system that aims to automate the recycling sorting process, thereby increasing the efficacy of current waste management systems and eliminating the risk to human lives. After applying Convolutional Neural Network, Decision Tree, and Random Forest classification models, we have designed and implemented an automated system that can identify and sort various waste items into trash and recyclable categories. The Eco-Bin aims to advance society toward self-sustainability by making environmentally friendly habits like recycling more convenient, safer, and cost-effective.

2 Background

2.1 Convolutional Neural Networks

Neural networks form the foundation of deep learning, and they are applied to a wide range of tasks, including forecasting and language processing. Convolutional Neural Networks (CNNs) are particularly renowned as highly effective deep learning models, especially well-suited for tasks in computer vision and classification [LeCun et al., 2015] which is why they are employed for this project's waste separation task. CNNs are constructed using three key types of layers: convolutional layer, pooling layer, and fully-connected layer [Yamashita et al., 2018]. The Convolutional Layer is the unique layer of CNN that applies a set of filters across an input image grid to create a feature map that contains information like edges, textures, and shapes. Using the feature map produced by the Convolutional Layer, the Pooling Layer simplifies the feature map to reduce computational costs and prevent model overfitting. The Fully Connected layer is the final layer responsible for connecting the details in the feature map to make a classification prediction. These layers are all linked through various weights and edges which are adjusted over various epochs during the training process to minimize loss.

2.2 Learning Algorithms

2.2.1 Decision Tree

Decision trees are simplified supervised machine-learning strategies that are frequently used in regression or classification tasks [Wu et al., 2008]. The learning framework begins by aggregating all training data into an initial root node and then recursively partitioning data into subsets across various features. Thus at each split into leaf nodes, the data becomes more classified until a terminating condition is met and all data is categorized. Due to their simplicity, Decision Trees are often susceptible to overfitting, variations in training data, and bias toward dominant classes. These limitations are mitigated by their extension algorithm, Random Forest.

2.2.2 Random Forest

Random Forest is an ensemble machine learning algorithm that combines multiple decision trees for more accurate results [Ho, 1995]. Random Forest uses a technique called Bagging to create multiple subsets of data through random sampling. Through Random Feature selection, the features are then sampled to create a subset of identifiable features. These random data sets and sets of features are then used to create multiple decision trees. Once all the decision trees are created, each tree identifies the class with the most shared data and features. These class identifications are then averaged across the forest for final predictions. By incorporating features from multiple trees, Random Forest is less susceptible to overfitting and noise.

3 Implementation

The designed Eco-Bin system has five key steps that it uses to automate the recycling process:

1. A waste item is first inserted inside the physical Eco-Bin

2. A camera is then used to take a picture of the waste

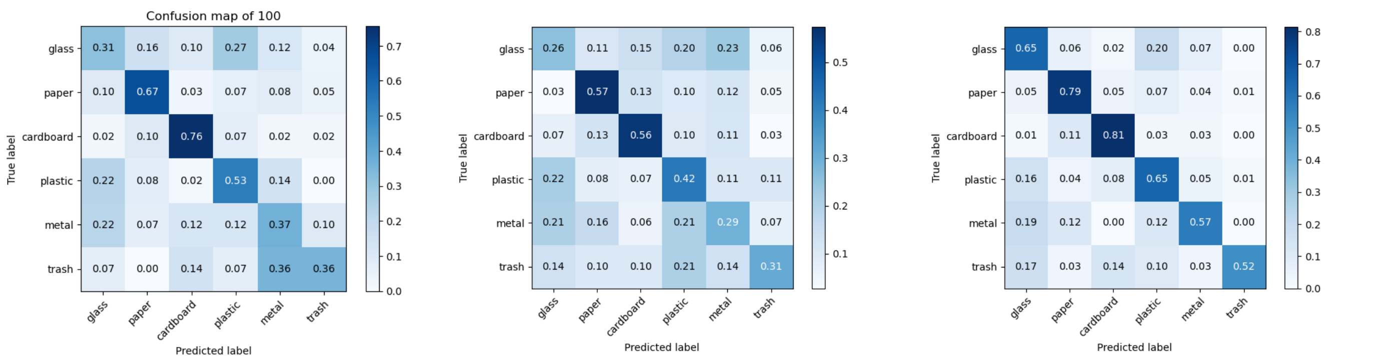
3. A classification model (CNN, Decision Tree, Random Forest) is used to classify the waste type

4. Based on the classification, the physical system then sorts the waste into disposable or recyclable categories

In summary, the key components of the system are the classification model and the physical bin. To test and train the Eco-Bin's CNN, we used the Kaggle Garbage Classification Data Set. This data set has 2527 images uniformly classified as cardboard, glass, metal, paper, or trash [Cchangcs, 2018]. Note that the images lack consistent blank space, zoom, or orientation simulating the environment of trash disposable within a bin. This data was then preprocessed using the sklearn library [Pedregosa et al., 2011]. For the Decision Tree and Random Forest Architectures, the images were first converted to greyscale, then resized to 32×32 pixels, and then flattened. For the CNN, the images were loaded from the camera, resized to 128×128 images, and then split into a $9 : 1$ ratio between the training and testing data sets. The model was created from 3 128×128 convolutional layers, 128 128×128 pooling layers, and 128 42×42 fully connected layers and trained over 8 epochs.

4 Results

Across the various classification models, we generated various confusion matrices to represent the model's accuracy when identifying materials and trash. These images are shown in figures $1, 2$, and 3 below: As



(a) Confusion Matrix of CNN (b) Confusion Matrix of Decision Tree (c) Confusion Matrix of Random Forest

Figure 1: Confusion Matrices of the Classification Models

shown the Random Forest classifier performed the best able to correctly identify the correct label about 75 percent of the time. Across the board, all classification models did pretty well in identifying paper, cardboard, and plastic, however glass, metal, and trash remained a challenge for all but the Random Forest Model.

5 Conclusion

Eco-Bin could revolutionize future waste management. The current MVPs are already able to distinguish between recyclable and trash materials with similar accuracy to most humans. Further work has to be done to improve accuracy further and test its accuracy on more complex types of waste, but current prototypes show potential. Overall Eco-Bin has the potential to make green sustainability more convenient.

References

[Adedeji and Wang, 2019] Adedeji, O. and Wang, Z. (2019). Intelligent waste classification system using deep learning convolutional neural network. *Procedia Manufacturing*, 35:607–612. The 2nd International Conference on Sustainable Materials Processing and Manufacturing, SMPM 2019, 8-10 March 2019, Sun City, South Africa.

[Cchangcs, 2018] Cchangcs (2018). Garbage classification.

[Ho, 1995] Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE.

[Lahcen et al., 2022] Lahcen, G., Mohamed, E., Mohammed, G., Hanaa, H., and Abdelmoula, A. (2022). Waste solid management using machine learning approch. In *2022 8th International Conference on Optimization and Applications (ICOA)*, pages 1–5.

[LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

[Nnamoko et al., 2022] Nnamoko, N., Barrowclough, J., and Procter, J. (2022). Solid waste image classification using deep convolutional neural network. *Infrastructures*, 7(4).

[Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[Poudel and Poudyal, 2023] Poudel, S. and Poudyal, P. (2023). Classification of waste materials using cnn based on transfer learning. In *Proceedings of the 14th Annual Meeting of the Forum for Information Retrieval Evaluation*, FIRE '22, page 29–33, New York, NY, USA. Association for Computing Machinery.

[Wu et al., 2008] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Philip, S. Y., et al. (2008). Top 10 algorithms in data mining. *Knowledge and information systems*, 14(1):1–37.

[Yamashita et al., 2018] Yamashita, R., Nishio, M., Do, R. K., and Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, 9(4):611–629.