Proceedings of the 5th International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2026.36577

Evaluating a ResNet-50-Based Interactive Image Classifier for Recycling Materials

Chenxi Huang

Crean Lutheran High School, Irvine, USA
chenxih208@gmail.com

Abstract. Accurate recycling relies on correctly identifying recyclables, but superficially identical things might make hand sorting ineffective. This work created and tested an image classifier based on an ImageNet-pretrained ResNet-50 for eleven types of recycling materials. A public dataset of 3,107 photos was split into 2,485 training and 622 test images. Transfer learning was utilized to tailor the pretrained network to the recycling categories. The classifier scored roughly 75% accuracy on the independent test set. The confusion matrix showed that the most common errors were between visually comparable categories, such as paper and cardboard and soft and hard plastics. The trained classifier was also used in a Gradio interface that takes an uploaded image and predicted category along with confidence values. The findings indicate that image classification may help with recycling education and preliminary sorting, but unclear predictions should be confirmed by a user and the system tested in real-world settings before widespread adoption.

Keywords: recycling materials, image classification, ResNet-50, transfer Learning, waste Management

1. Introduction

Recycling minimizes landfill space, conserves resources and contributes to a circular economy. Municipal solid trash is increasing, and the need for precise sorting is growing since contaminants can lower the value and usefulness of collected items. The U.S. Environmental Protection Agency (EPA) has adopted a national objective of 50% recycling by 2030 [1]. EPA's Waste Reduction Model also allows users to analyze greenhouse-gas impacts of recycling, landfilling and other materials-management activities [2]. The environmental stakes drive the development of accessible technologies that can help consumers identify typical recycling materials.

With recent developments in computer vision, image classification has become a feasible method for waste recognition. Widely available software can give consistent forecasts using a camera-based approach. This paper explores the question: How good is a pretrained ResNet-50 at classifying eleven recycling-material categories, and what do its error patterns mean for real-world recycling support? The scope is limited to image classification with one public dataset and no robotic sorting, more sensors and conveyor-belt automation. This study examines the overall performance and the performance at the class level and indicates areas where more data, better definitions of categories, or human review might be necessary.

DOI: 10.54254/2754-1169/2026.36577

2. Literature review

Convolutional neural networks (CNNs) learn hierarchical visual features directly from image data and form the basis of modern image classification [3]. AlexNet helped establish deep CNNs as a leading approach to large-scale image recognition [3]. Later ResNet introduced residual connections which made training much deeper networks easier and reduced degradation as depth increased [4]. Transfer learning allows a model pretrained on a large dataset such as ImageNet to be adapted to a smaller task-specific dataset [5]. For this study, an ImageNet-pretrained ResNet-50 offered a practical starting point because the available recycling dataset was moderate in size.

Waste-image research has moved from controlled datasets toward more complex real-world settings. Thielmann et al. organize the field into controlled datasets, in-the-wild images, and industrial conveyor-belt systems, emphasizing that performance depends strongly on the deployment environment [6]. TrashNet introduced a six-category benchmark of individual waste objects photographed against clean backgrounds [7]. TACO later provided litter images from varied outdoor settings for detection and segmentation [8]. Nnamoko et al. described a complete end-to-end solid-waste classification workflow including preprocessing, evaluation, computational requirements and limitations [9]. Together, these studies show why accuracy values from different datasets are not directly comparable. The present study therefore interprets its results within its own eleven-category dataset.

Waste classification also has environmental and practical significance. EPA's national goal calls for increasing the U.S. recycling rate to 50% by 2030 [1], and its Waste Reduction Model estimates the greenhouse-gas consequences of different materials-management practices [2]. Consumer-facing applications have also been explored. Yong et al. combined MobileNetV2 with a WeChat mini-program that allowed users to upload an image and receive a waste category [10]. However, much of the literature emphasizes industrial systems, benchmark datasets, or lightweight mobile classifiers. This study instead combines a pretrained ResNet-50 with a simple Gradio interface and evaluates it on an eleven-category recycling dataset.

The choice of dataset changes the meaning of a reported accuracy. TrashNet uses single objects on clean backgrounds, so the main problem is separating material categories when the object is easy to see [7]. TACO is different. Its images contain litter in outdoor scenes, where objects may be small, partly hidden, or surrounded by unrelated visual information [8]. Industrial systems add another layer of difficulty because materials move quickly and may overlap on a conveyor belt [6]. These settings are connected, but they are not interchangeable. A model that performs well on one of them may not perform equally well on another.

The earlier studies also differ in what they ask the model to do. TrashNet focuses on image-level classification, whereas TACO supports detection and segmentation in complex scenes [7, 8]. Nnamoko et al. provide a useful comparison because their work describes preprocessing, model evaluation, hardware requirements, and experimental limits as parts of one complete workflow [9]. Yong et al. move in a more user-facing direction by connecting MobileNetV2 to a WeChat mini-program [10]. That study is especially relevant to the present project. It shows that access and interface design matter alongside accuracy. The current study follows a similar practical goal, but uses ResNet-50 and a browser-based Gradio interface.

DOI: 10.54254/2754-1169/2026.36577

3. Methodology

3.1. Overall workflow

Figure 1 summarizes the study workflow. Images were collected, reviewed, preprocessed, and divided into training and test sets. An ImageNet-pretrained ResNet-50 was then fine-tuned to classify eleven recycling-material categories. Performance was evaluated on the independent test set using overall accuracy and a confusion matrix. Finally, the trained classifier was deployed in a Gradio interface that accepts an uploaded image and returns a predicted category with confidence scores.

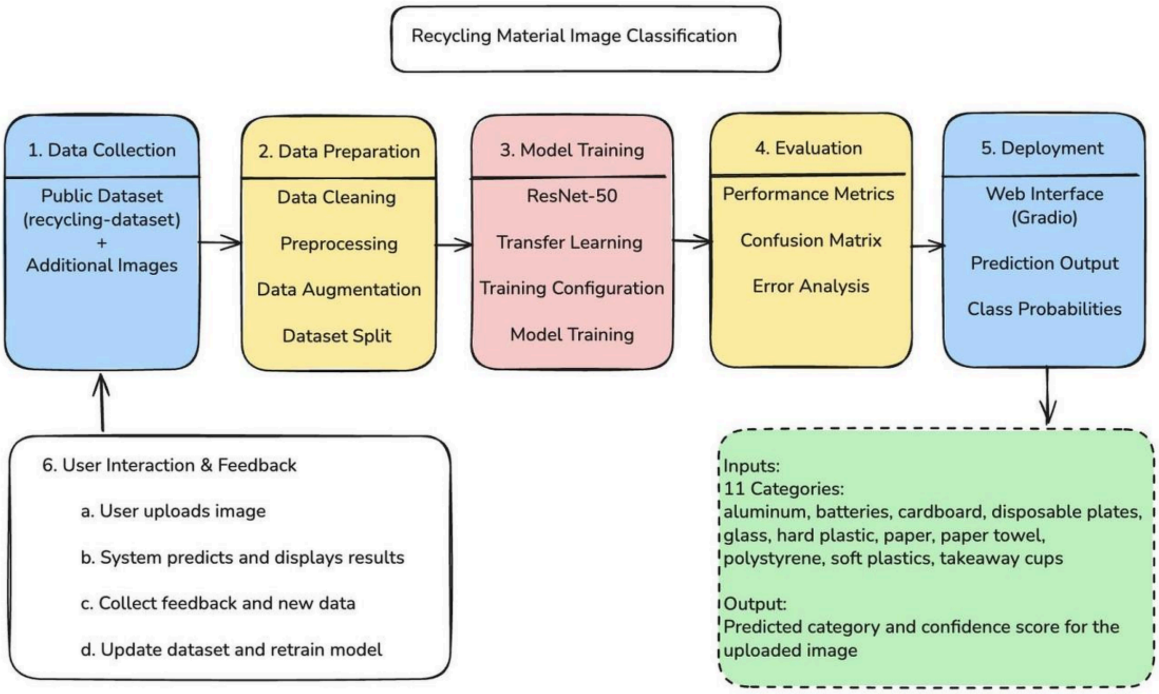

Figure 1. Overview of the solution

3.2. Dataset collection

The main source was the public viola77data Recycling Dataset hosted on Hugging Face. The dataset contains 3,107 labeled images in eleven categories: aluminium, batteries, cardboard, disposable plates, glass, hard plastic, paper, paper towel, polystyrene, soft plastics, and takeaway cups.

Figure 2 shows sample images covering all 11 groups of materials: aluminum, batteries, cardboard, disposable plates, glass, hard plastic, paper, paper towels, polystyrene, soft plastic and takeaway cups.

The complete data set includes 3,107 pictures. In all samples, 2,485 pictures constitute the training set, and 622 pictures constitute the test set, accounting for 80% and 20% of the whole data set respectively. During the whole training process, the test set was kept separate and only used for the final model evaluation.

DOI: 10.54254/2754-1169/2026.36577

Figure 2. Representative examples of 11 categories of recycled materials used in the data set

3.3. Data preparation

Before model training, all images were checked for label quality and usability. Each image was adjusted to match the input size standard required by ResNet-50. Normalization was carried out according to the standard preprocessing rules used for pre-trained models on the ImageNet data set.

Data augmentation technology was adopted in the training stage. These methods can make reasonable changes to the image without changing its original category label. This practice helps to enhance the generalization ability of the model and reduce the risk of overfitting.

3.4. Classifier development

Model training relies on the Hugging Face framework and runs on an NVIDIA T4 GPU. The classifier adopts AdamW optimizer and cross-entropy loss. The training parameters are set to the learning rate of 0.001, with a total of 15 training epochs.

Overfitting control is an important problem in this experiment. Without proper restrictions, the model may only remember the training images instead of learning the general visual features. This work uses the ResNet-50 model pre-trained on ImageNet. The existing visual knowledge of the pre-trained model has been adjusted to adapt to the task of recycling material identification, and data augmentation is applied as auxiliary support.

3.5. Classifier evaluation

After the training, the classifier was evaluated using a separate test set. The overall classification accuracy, calculated according to the proportion of test images assigned to the correct material group, serves as the main evaluation indicator.

A confusion matrix was created to analyze the prediction results of all 11 categories. Category-level observation may expose repeated misclassification cases, which cannot be reflected solely by overall accuracy. The matrix also shows which types of materials pose the greatest identification challenges.

DOI: 10.54254/2754-1169/2026.36577

The results of error analysis support several feasible optimization plans: collecting additional images for categories that are difficult to distinguish, defining material categories more accurately, optimizing preprocessing steps, and adding manual verification when the confidence of model output prediction is low.

4. Results

4.1. Overall performance

The ResNet-50 model with pre-trained weights achieved an accuracy of about 75% on the independent test set. In other words, the model correctly identifies most of the 622 test images belonging to 11 recycling categories. Since these test images do not appear during training, the accuracy value reflects the performance of the model on new samples from the same data set.

The confusion matrix generated from the test set results can be seen in Figure 3. Most of the prediction results fall on the diagonal, which corresponds to the correct classification. In this test group, there are relatively few mispredictions generated by batteries, paper towels, disposable plates and takeaway cups. This shows that the classifier is easier to identify the visual features of these items.

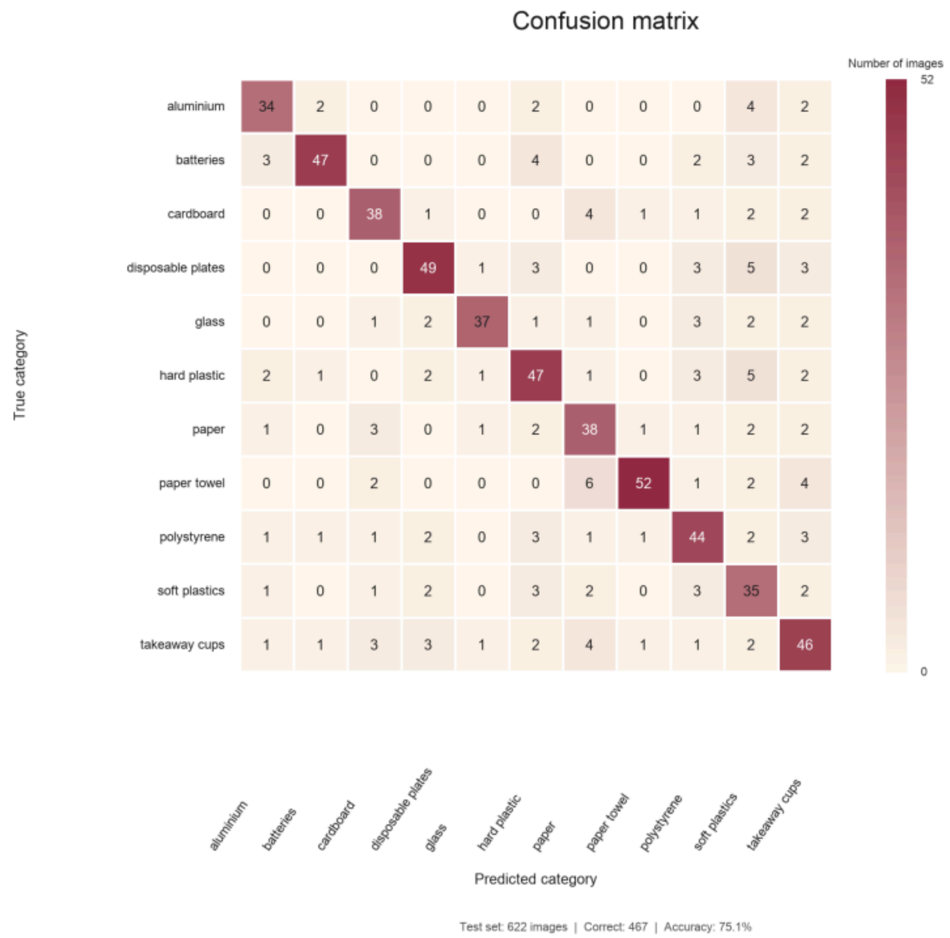

Figure 3. Confusion matrix of the ResNet-50 classifier on the testing dataset

DOI: 10.54254/2754-1169/2026.36577

4.2. Error analysis

The confusion matrix shows frequent misclassifications between visually similar categories. Several paper samples were mistakenly marked as cardboard. Some soft plastic images were classified as hard plastic. A small number of disposable plate and polystyrene samples were also mistakenly identified as soft plastics. Such errors occur mainly due to the common characteristics of color, shape and surface texture.

These results prove that visual similarity is the main reason for classification errors. Differences in lighting conditions, shooting angles and background environment may also make it more difficult to identify some pictures. Although these misjudgments will reduce overall accuracy, they provide clear guidance on which categories need more diverse training pictures.

Therefore, Figure 3 provides additional information that cannot be provided by overall accuracy alone. It marks where classification errors occur and identifies the categories that are most likely to benefit from extended data sets, clearer class definitions, or manual checks.

Table 1. The most common misclassification patterns observed in the test data set

True Category	Incorrect Prediction	Number of Images
Paper	Cardboard	3
Soft plastics	Hard plastic	3
Disposable plates	Soft plastics	2

Table 1 lists the most common error patterns observed from the confusion matrix. Three paper images were mistakenly classified as cardboard, and three soft plastic images were identified as hard plastic. Two disposable plate images and two polystyrene images were mistakenly classified as soft plastic. Based on these results, targeted improvements can be made: collect more samples for these problem categories and clarify the subtle category boundaries where visual differences exist.

Figure 4 shows how the user interface handles uncertain predictions. It displays the first three predicted confidence values and waits for the user's confirmation. The percentage value displayed is for demonstration purposes only. If accurate experimental results are needed, these numbers should be replaced by the real output of the trained model.

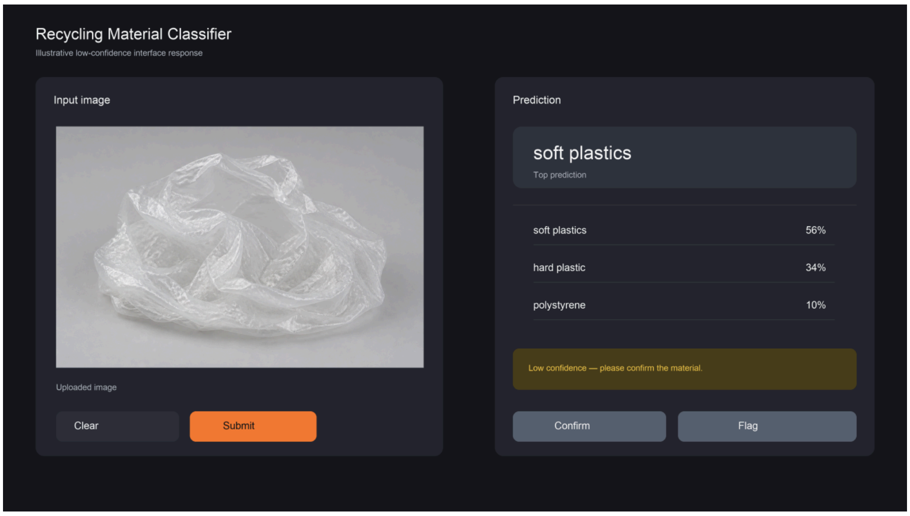

Figure 4. Low confidence output illustration in the Gradio interface

DOI: 10.54254/2754-1169/2026.36577

5. Discussion

When using unseen test data for evaluation, the accuracy of the ResNet-50 classifier is close to 75%. This result shows that transfer learning can effectively handle this medium-scale data set, but the difference in classification performance between different material groups is obvious. Items with unique visual characteristics are usually reliably identified, and misidentification often occurs between materials with similar appearances. Paper samples are occasionally marked as cardboard, and soft plastic is sometimes mistaken for hard plastic. This repeated trend shows that visual similarity is the main obstacle affecting the performance of the classifier.

These recurring error patterns also point to the actual direction of further optimization. Paper and cardboard usually have comparable tones and surface textures, while hard plastic and soft plastic may look indistinguishable in a static photo. In future iterations, the system interface can display the highest predicted confidence value and prompt the user to confirm the material type when the model produces an uncertain output. Expanding the training collection by adding images captured under different lighting conditions, different backgrounds and multiple angles can also improve the recognition accuracy of these easily confused categories.

It should be clarified that the prototype developed in this work serves as a network-based lightweight decision-making aid, not a comprehensive industrial classification system. According to existing records, AMP operates three large facilities and deploys more than 400 artificial intelligence sorting units in North America, Asia and Europe [11]. A report released by the Ellen MacArthur Foundation points out that AMP Cortex can handle up to 80 objects per minute with an accuracy rate of 99% [12]. This kind of system relies on dedicated robot hardware to handle high-throughput waste separation. In contrast, the Gradio prototype built for this study only allows users to upload a single image through a standard web browser and receive predicted categories and corresponding confidence scores. Therefore, the tool can be used as a useful supplement to industrial sorting equipment and has potential applications in homes, schools, university campuses and local community recycling programs. However, before drawing a solid conclusion on its practical effectiveness in such situations, real-world deployment tests are required.

These results also have a wider environmental relevance. As shown in Figure 3, the vast majority of repeated misclassifications only come from a few pairs of visually similar materials. Therefore, requiring users to verify uncertain predictions can improve source-level sorting judgements. A West Coast survey on recycling contamination recorded an average contamination rate of about 11% in Oregon and Washington, while the average contamination rate in California was 20%; some California communities even faced contamination levels of up to 46% [13]. These statistics show that making better classification decisions before waste collection may help reduce contamination rates, although field tests and user evaluations are necessary to confirm the actual benefits of the prototype.

Several limitations of this study are worthy of acknowledgement. The experiment only covers 11 material categories and relies on an open image library and a browser-operated prototype. All tests were completed in a controlled experimental environment, not at active real-world recycling stations. Subsequent research may introduce additional material categories, incorporate more diverse photos taken under daily conditions, and explore object detection, video recognition and mobile platform adaptation. Field research is crucial to confirm whether this classification method can actually improve recycling-related decision-making.

There is an obvious trend in the whole error analysis process: classification errors are unevenly distributed across all 11 categories. The most common misjudgment occurs between materials that look similar in the photo. Paper and cardboard usually have similar colors and textures. In addition, when the camera angle is limited and it is difficult to judge the thickness and shape, thin and rigid plastic containers may be confused with soft plastic products. This limitation occurs because a single image cannot reflect physical characteristics such as weight and flexibility, allowing the classifier to make judgments purely based on visible visual characteristics.

Figure 4 shows a feasible solution to this weakness. The suggested interface will not only display a single prediction label, but will display multiple confidence scores once there is uncertainty in the prediction and user confirmation is sought. Under this design, the classifier operates as a supporting device instead of providing definitive results. This arrangement is more realistic for students, home users and campus recycling programs, rather than assuming that every automatic prediction is correct. In addition, these challenging samples can be marked for future checks and may be used to expand and enhance training datasets.

6. Conclusion

This study evaluated the pre-trained ResNet-50 model to classify 11 types of recyclable materials. Based on a public data set consisting of 3,107 pictures and transfer learning strategies, the classifier achieves an accuracy rate of about 75% on the independent test set. The information derived from the confusion matrix shows that the visual similarity between different materials explains most of the classification errors.

The study contributes an evaluation of a ResNet-50-based recycling-material classifier and an error analysis that points to specific opportunities for improvement. Deploying the trained classifier through a Gradio interface also demonstrates how deep learning can be translated into an accessible recycling-support tool for schools, campuses, communities, and households.

This study has certain drawbacks. The dataset was limited and the evaluation was done in controlled conditions rather than a real recycling setting. Future work should extend the dataset, test further model architectures and evaluate the system on a wider variety of real-world photos. User studies would also be needed to determine if the tool does improve recycling decisions.

References

[1] U.S. Environmental Protection Agency. (n.d.). U.S. national recycling goal. *U.S. Environmental Protection Agency*. https://www.epa.gov/circulareconomy/us-national-recycling-goal

[2] U.S. Environmental Protection Agency. (2024). Waste reduction model (WARM). *U.S. Environmental Protection Agency*. https://www.epa.gov/waste-reduction-model

[3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional networks. *Communications of the ACM, 60*(6), 84–90. https://dl.acm.org/doi/10.1145/3065386

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

[5] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering, 22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

[6] Thielmann, P., Zhou, Y., Mirbach, B., Stricker, D., & Rambach, J. R. (2025). A review of computer vision for industrial-grade waste classification. *IEEE Access, 13*, 151934–151953. https://doi.org/10.1109/ACCESS.2025.3603903

[7] Yang, M., & Thung, G. (2016). Classification of trash for recyclability status. *CS229 Project, Stanford University*. https://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf

[8] Proença, P., & Simões, J. (2020). TACO: Trash annotations in context for litter detection. *arXiv preprint arXiv:2003.06975*. https://arxiv.org/abs/2003.06975

[9] Nnamoko, N., Barrowclough, J., & Procter, J. (2022). Solid waste image classification using deep convolutional neural network. *Infrastructures, 7*(4), 47. https://doi.org/10.3390/infrastructures7040047

DOI: 10.54254/2754-1169/2026.36577

[10] Yong, L., Ma, L., Sun, D., & Du, L. (2023). Application of MobileNetV2 to waste classification. *PLOS ONE, 18*(3), e0282336. https://doi.org/10.1371/journal.pone.0282336
[11] AMP Robotics Corp. (2025). AMP named to 2025 global cleantech 100. *AMP Robotics*. https://ampsortation.com/articles/amp-named-2025-cleantech-100
[12] Ellen MacArthur Foundation. (2021). Artificial intelligence for recycling: AMP robotics. *Ellen MacArthur Foundation*. https://www.ellenmacarthurfoundation.org/circular-examples/artificial-intelligence-for-recycling-amp-robotics
[13] The Recycling Partnership. (2020). 2019 West Coast contamination initiative research report. *The Recycling Partnership*. https://recyclingpartnership.org/wp-content/uploads/dlm_uploads/2020/04/West-Coast-Contamination-Initiative-Report-6.22.20.pdf