# A Study of Bayesian and Convolutional Models for Waste Classification

**Mason Kulikowski**
University of Delaware
masonkul@udel.edu

## Abstract

Automated waste detection is a crucial component in enhancing recycling efficiency. This study compares the performance of a feature-based Bayesian classifier and a convolutional neural network (CNN) in multi-class classification. Both models are evaluated on the TrashNet dataset, tasked with classifying images of recycling and trash into six categories: glass, paper, plastic, cardboard, metal, and trash. The Bayesian classifier relies on engineered color, texture, and shape features, while the CNN directly learns hierarchical representations from image data. Experimental results show comparable overall performance between the two approaches. However, the CNN exhibits significant difficulty in correctly identifying the trash class due to class imbalance, frequently misclassifying trash samples as recyclable materials. The observed misclassification behavior emphasizes the role of class imbalance in CNN-based waste classification.

## Introduction

Accurate waste sorting is essential for effective recycling systems, as contamination from improperly classified materials can significantly reduce recycling efficiency. Many recycling facilities rely on manual sorting and multi-stage processing pipelines that are time-intensive and resource-demanding, often resulting in inconsistent sorting quality. Visual similarities between recyclable materials and trash further complicate the sorting process, making reliable classification difficult at scale. To address these challenges, this study investigates an image-based machine learning approach for automatically distinguishing recyclable materials from trash. The objective of this work is to evaluate the effectiveness of automated visual classification methods in a controlled setting and assess their potential to reduce labor demands while improving operational efficiency in recycling facilities.

To investigate this problem, this study adopts a comparative modeling framework that evaluates two fundamentally different approaches to image-based waste classification. A feature-based Bayesian classifier and a CNN are implemented and benchmarked on the same classification task to ensure a consistent basis for comparison. The Bayesian approach relies on manually engineered visual descriptors that encode domain-relevant information such as color, texture, and shape, enabling transparent modeling assumptions and lower computational complexity. In contrast, the CNN is designed to learn hierarchical feature representations directly from raw image data, allowing it to capture complex visual patterns without explicit feature engineering. By analyzing both models under comparable experimental conditions, this work aims to assess how model design choices influence classification behavior, robustness across material classes, and sensitivity to data characteristics commonly encountered in recycling applications.

## Related Work

Prior work in image-based waste classification has focused on applying computer vision techniques to distinguish recyclable materials from trash. Thung and Yang (2016) [1] introduce the TrashNet dataset and formulate waste classification as a multi-class image recognition problem involving common material categories. Their study evaluates both traditional machine learning methods, including a support vector machine (SVM) trained on hand-crafted visual features, and convolutional neural network-based approaches. This work demonstrates the feasibility of automated waste classification and provides a dataset that serves as the foundation for the experiment conducted in this study.

The convolutional neural network architecture used in this study is informed by design principles established through large-scale image classification research, particularly work conducted using the ImageNet dataset. ImageNet enabled systematic evaluation of deep convolutional neural networks for hierarchical feature learning and influenced common architectural patterns such as stacked convolutional layers, progressive spatial downsampling, and nonlinear activations [2]. Although ImageNet addresses a general object recognition task, the architectural conventions and benchmarking practices derived from its use informed both the design and the evaluation methodology applied in this project.

More recent studies have focused on improving waste classification performance through model optimization and system-level design. Nahiduzzaman et al. propose an optimized waste classification pipeline that combines lightweight convolutional networks with ensemble-based classifiers to improve accuracy and computational efficiency [3]. Their approach emphasizes architectural refinement and comparative evaluation against baseline models. In contrast, this work does not pursue performance optimization; instead, it compares a feature-based probabilistic classifier and a CNN under consistent experimental conditions to analyze differences in model behavior and sensitivity to dataset characteristics.

## Method

### Bayesian Classifier

A Bayesian classifier operates by calculating posterior probabilities for each class given an observed feature vector and assigning the class with the highest posterior probability. This is achieved by modeling the likelihoods of the extracted features and combining them with prior class probabilities using Bayes' theorem:

$$P(y = k \mid x) = \frac{P(x \mid y = k)\, P(y = k)}{P(x)}$$

Where $x$ denotes the extracted feature vector and $y = k$ represents the predicted class label. Since $P(x)$ is constant across classes, classification is done by selecting the class that maximizes $P(x \mid y = k)\, P(y = k)$.

A structured feature extraction pipeline was employed to map each input image to a feature vector for Bayesian classification. The feature set encodes visual characteristics relevant to material discrimination, including color distribution, surface texture, shape properties, and material-specific reflectance cues. The design emphasizes interpretability and supports probabilistic inference without relying on learned representations.

Color and texture features were used to capture surface-level visual differences across materials. Images were converted from RGB to the HSV color space, where hue represents dominant color, saturation represents color intensity, and value represents brightness. For each channel, the mean and standard deviation were computed to describe the average color and how much it varies across the image. Texture information was encoded using Local Binary Patterns (LBP) and Gray-Level Co-occurrence Matrix (GLCM) statistics. LBP captures fine-scale texture by encoding local intensity differences between neighboring pixels, while GLCM statistics describe broader spatial relationships between pixel intensities. Haralick-style properties, including contrast, energy, and homogeneity, were used to represent differences in surface regularity and texture consistency. Together, these features help distinguish materials with uniform surfaces, such as paper and glass, from rougher or more heterogeneous materials, such as cardboard and trash.

Shape and material-specific features were incorporated to capture high-level visual cues not fully represented by color or texture alone. Global shape information was encoded using image aspect ratio and Hu moments, which provide rotation and scale-invariant descriptors derived from image moments. These features capture coarse geometric structure and help differentiate materials with distinct object profiles. In addition, reflectance properties were explicitly modeled through gradient-based features. Specular reflection measures were used to quantify the presence and concentration of sharp highlights, which are characteristic of reflective materials such as metal and glass. Directional edge patterns were included to capture structured reflections commonly observed on metallic surfaces, while brightness smoothness and saturation uniformity were used to represent visual properties associated with glass. A texture variability feature was also included to model the heterogeneous appearance of trash, which lacked a consistent visual structure in the dataset.

Images often contain correlated features, as multiple features may encode related visual properties derived from the same underlying pixel information. Such correlations are common among texture, reflectance, and multi-scale features and violate the conditional independence assumption in Bayesian classifiers, biasing likelihood estimates. To address this, a correlation matrix was computed to quantify pairwise feature dependencies, and highly correlated features were removed based on a threshold of 0.85. Following pruning, principal component analysis (PCA) was applied to further reduce dimensionality and correlation by projecting the feature space onto an orthogonal basis that preserves the majority of variance. This process produces a compact, decorrelated feature representation that improves stability and is better aligned with the assumptions of Bayesian inference.

### Convolutional Neural Network

The constructed CNN closely models the ImageNet and TrashNet models. Different convolutional and pooling blocks were implemented to capture varying levels of information in the images. A 15-layer CNN was constructed that is structurally similar to the models mentioned above.

- **Layer 0:** Input image of size 224x224
- **Layer 1:** Convolution with 32 filters, size 3x3, stride 1, padding 1
- **Layer 2:** Convolution with 32 filters, size 3x3, stride 1, padding 1
- **Layer 3:** Max-Pooling with size 2x2, stride 2
- **Layer 4:** Convolution with 64 filters, size 3x3, stride 1, padding 1
- **Layer 5:** Convolution with 64 filters, size 3x3, stride 1, padding 1
- **Layer 6:** Max-Pooling with size 2x2, stride 2
- **Layer 7:** Convolution with 128 filters, size 3x3, stride 1, padding 1
- **Layer 8:** Convolution with 128 filters, size 3x3, stride 1, padding 1
- **Layer 9:** Max-Pooling with size 2x2, stride 2
- **Layer 10:** Convolution with 256 filters, size 3x3, stride 1, padding 1
- **Layer 11:** Convolution with 256 filters, size 3x3, stride 1, padding 1
- **Layer 12:** Max-Pooling with size 2x2, stride 2
- **Layer 13:** Global Average Pooling
- **Layer 14:** Fully Connected layer with 128 neurons
- **Layer 15:** Fully Connected layer with 6 neurons
- **Result:** Non-normalized class logits for 6 waste categories

## Experiment

### Dataset

All experiments were conducted using the TrashNet dataset, a publicly available image collection for automated waste classification. The dataset contains 5,054 color images captured under controlled lighting with a uniform background, minimizing environmental noise. Images vary in shape, texture, reflectance, and material appearance, introducing meaningful visual variability. The dataset also has an uneven distribution of samples across categories, which was intentionally preserved to reflect realistic data conditions and to examine how different models respond to class imbalance and visual variability.

Fig 1. Glass

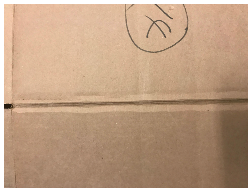

Fig 2. Cardboard

Fig 3. Metal

Fig 4. Paper

Fig 5. Plastic

Fig 6. trash

Data augmentation was applied during training to increase the effective dataset size and improve robustness. Augmentations included random rotations and color jittering to introduce controlled variation. Size normalization was used to ensure architectural consistency for the CNN.

**Bayesian Classifier**

For the Bayesian model, extracted feature vectors were used to train a Gaussian Naive Bayes classifier using the training split of the dataset. Empirical class priors were estimated from the training data, and the trained model was evaluated on a held-out test set. The same data split and evaluation metrics were used as for the CNN to ensure a consistent basis for comparison.

**Convolutional Neural Network**

The CNN was trained for 50 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 16. A small weight decay (0.0001) was applied for regularization. Learning rate scheduling was enabled using a step scheduler, reducing the learning rate by a factor of 0.1 every 10 epochs. Early stopping was employed with a patience of 10 epochs to prevent overfitting and ensure stable training.

**Results**

Model performance was evaluated using precision and recall, which capture false-positive and false-negative rates produced by the models. These metrics provide a clearer view of class-specific prediction behavior than overall accuracy, particularly for visually heterogeneous samples.

**Bayesian Classifier**

A 70/30 testing and training split was used for the Bayesian model; it achieved a test accuracy of 59.95%. This slump in accuracy is related to the Bayesian assumption of independence and the features violating this assumption.

| Material | Precision | Recall |
|---|---|---|
| glass | 0.5120 | 0.5519 |
| paper | 0.7198 | 0.7278 |
| cardboard | 0.8557 | 0.7155 |
| metal | 0.4833 | 0.4874 |
| plastic | 0.6720 | 0.5563 |
| trash | 0.2029 | 0.3590 |

That doesn't mean the model didn't perform well on certain classes. Bayes performed well on classes with solid structures like cardboard and paper, but struggled with classes that were transparent like glass and plastic. Many of the trash examples were also transparent and created more confusion for the classifier.

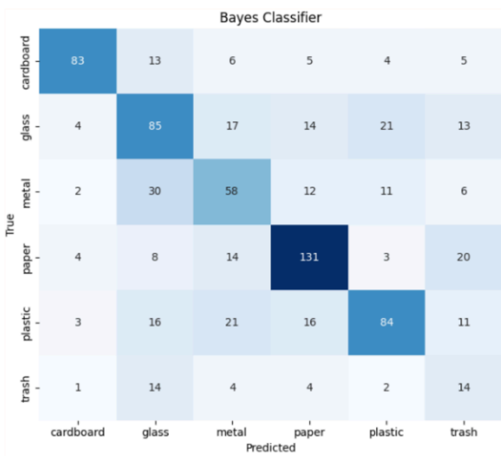


Fig. 7 Confusion matrix for Bayes

**Convolutional Neural Network**

CNN used a train/val/test split of 70/15/15 and achieved slightly higher test accuracy than Bayes with 60.87%. The interesting thing is that looking at the precision and recall metrics,

| Material | Precision | Recall |
|---|---|---|
| glass | 0.3667 | 0.5000 |
| paper | 0.7486 | 0.7611 |
| cardboard | 0.8214 | 0.7931 |
| metal | 0.5888 | 0.5294 |
| plastic | 0.6327 | 0.6159 |
| trash | 0.0000 | 0.0000 |

CNN performed better on most classes except the trash class. The CNN failed to predict any test samples as trash and instead preferred trying to mislabel them as a form of recycling. Since there are so few trash samples, during optimization the model is biased towards the other classes compared to the trash class.

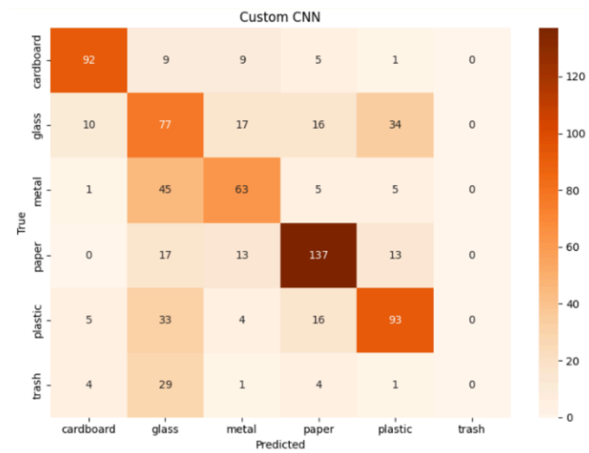


Fig. 8 Confusion matrix for CNN

Using balanced class samples would eliminate the zero trash problem, but wouldn't significantly increase the trash class because of the nature of trash samples. There are not consistent features to derive from trash and as a result models will tend to lean towards other recyclable classes.

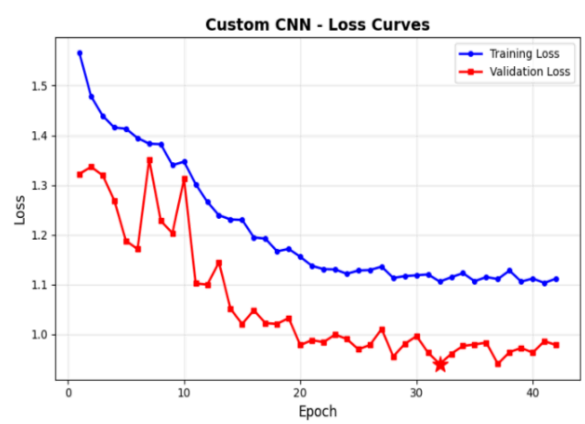


Fig. 9 Training and validation for CNN

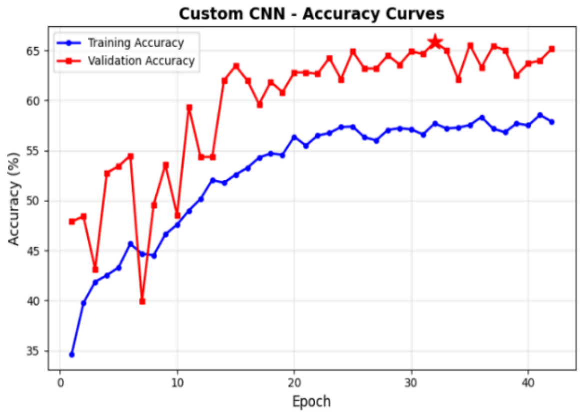


Fig. 10 Traing and validation accuracy for CNN

The CNN stopped training after 42 of 50 epochs when both training and validation accuracy plateaued and began to fluctuate. The loss curves continued to decrease without diverging, suggesting the model did not overfit but instead reached a learning limit. Additional optimization mainly increased confidence on majority recyclable classes rather than improving overall discrimination. This explains why the model failed to identify trash samples and highlights a limitation of CNN training under class imbalance and high visual variability.

## Conclusion

Machine learning shows promise for automated waste classification, but reliable performance is limited by visual ambiguity and the poorly defined nature of trash as a class. In this study, both models struggled with classifying trash, but exhibited different failure modes: the CNN failed to detect trash samples due to class imbalance and visual heterogeneity, while the Bayesian model demonstrated more stable behavior through explicit feature modeling. These results indicate that increased model complexity does not necessarily yield more reliable classification and that successful waste classification depends on aligning model assumptions with dataset characteristics.

## References


[1] M. Yang and G. Thung, "Classification of Trash for Recyclability Status," 2016. Available: https://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf

[2] L. Fei-Fei, J. Deng, and K. Li, "ImageNet: Constructing a large-scale image database," *Journal of Vision*, vol. 9, no. 8, pp. 1037–1037, Mar. 2010, doi: https://doi.org/10.1167/9.8.1037.

[3] Md. Nahiduzzaman et al., "An Automated Waste Classification System Using Deep Learning Techniques: Toward Efficient Waste Recycling and Environmental Sustainability," Knowledge-Based Systems, pp. 113028–113028, Jan. 2025, doi: https://doi.org/10.1016/j.knosys.2025.113028.