Data Science and Management Project: Garbage Classification Using Machine Learning

Authors: Benassati Andrea and Lodesani Simone

---

1. Introduction to the Problem

The exponential growth in waste production driven by rapid economic and industrial development has created significant environmental challenges worldwide. Effective waste management strategies are essential to mitigate environmental pollution, reduce resource depletion, and promote sustainable development. One of the most critical steps in waste management is the accurate classification and segregation of different types of waste materials, which enables proper recycling, reduces landfill burden, and facilitates resource recovery.

Traditional waste sorting relies heavily on manual labor, which is time-consuming, expensive, and prone to human error. Workers must physically inspect and sort waste items, leading to inconsistent classification accuracy and potential safety hazards from handling hazardous materials. These limitations have motivated the development of automated waste classification systems that leverage advances in computer vision and machine learning.

1.1 Problem Statement

This project addresses the challenge of automated garbage classification using image-based machine learning techniques. The primary objective is to develop and compare multiple classification models capable of distinguishing between different categories of waste objects based on photographic images. Specifically, the system must classify garbage into six distinct categories: **cardboard, glass, metal, paper, plastic, and trash**.

The problem is formulated as a **supervised multi-class classification task**, where each image is assigned to exactly one waste category. The dataset exhibits several challenging characteristics that make this a non-trivial problem:

- **Class imbalance**: Some waste categories (e.g., trash) are significantly underrepresented compared to others (e.g., paper), which can bias model predictions toward majority classes.
- **Visual variability**: Objects within the same category may have diverse appearances, colors, shapes, and textures, requiring models to learn robust feature representations.

- **Environmental factors**: Images may contain varying lighting conditions, backgrounds, and object orientations that affect classification performance.
- **Intra-class similarity**: Some waste types (e.g., paper and cardboard, or different types of plastic) may appear visually similar, requiring fine-grained discrimination.

1.2 Project Objectives

The primary objectives of this project are:

1. **Data management**: Implement a structured data storage solution using PostgreSQL database to manage image metadata, ensuring reproducibility and scalability.
2. **Exploratory analysis**: Perform comprehensive exploratory data analysis to understand dataset characteristics, class distributions, and potential preprocessing requirements.
3. **Model comparison**: Develop and evaluate multiple classification approaches with increasing complexity, from baseline models to deep learning architectures.
4. **Performance evaluation**: Assess model performance using appropriate metrics including accuracy, confusion matrices, and class-wise performance measures.
5. **Real-world testing**: Create a custom test dataset with photographs taken by the project team to evaluate model generalization to real-world conditions.

The project follows a progressive modeling strategy that aligns with data science best practices, starting from simple baseline models to establish performance benchmarks, then advancing to more sophisticated deep learning approaches that can capture complex visual patterns.

---

2. Related Work

Computer vision and machine learning for waste classification have advanced significantly in recent years. Early approaches used traditional machine learning with handcrafted features (HOG, color histograms, textures) fed into SVM or Random Forest classifiers. While some hybrid models achieved high accuracy, they often struggled with complex visual variations.

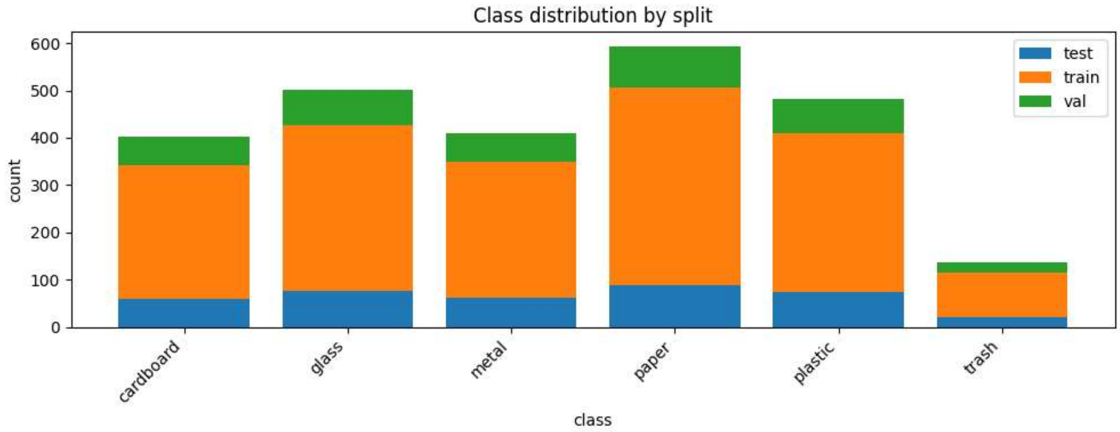
Convolutional Neural Networks (CNNs) revolutionized the field by automatically learning hierarchical features from raw pixels. Studies using CNN architectures achieved accuracies ranging from 60% to 99% depending on dataset complexity. Transfer learning further improved performance by leveraging pre-trained models (InceptionV3, ResNet50, MobileNet) fine-tuned for waste classification, particularly valuable with limited training data.

Recent innovations include data augmentation to improve generalization, model fusion combining multiple architectures, and integration with object detection systems for real-time robotic sorting applications. Our work implements a systematic comparison from baseline to transfer learning models, emphasizes structured data management with PostgreSQL, and augments limited training data with custom real-world photographs.

3. Methodology

3.1 Data Management

We implemented a hybrid storage system separating image files from metadata. Raw images (2,527 total) are stored locally while a PostgreSQL database manages metadata including file paths, labels, split assignments, and provenance. The dataset was consolidated and split stratified by class: 70% training (1,768 images), 15% validation (379), and 15% test (380). This addresses class imbalance where "trash" has only 96 training samples versus 416 for "paper".



3.2 Data Pipeline

Images are loaded on-the-fly during training to avoid memory constraints. The pipeline queries the database for metadata, creates TensorFlow datasets, and applies preprocessing: resizing to 192x256 (H x W) pixels (baseline CNN, MobileNetV2), normalization to [0,1] range (baseline CNN), re-standardization to [0,255] (MobileNetV2), and data augmentation (rotations, flips, zooms) applied only to training data.

3.3 Modeling Strategy

We implemented a progressive modeling approach with four models of increasing complexity:

1. **Logistic Regression (Baseline)**: Simple linear classifier operating on flattened, downsampled grayscale images (64×48 pixels) to establish a performance lower bound.
2. **Baseline CNN**: Custom architecture with three convolutional blocks (32, 64, 128 filters), batch normalization, max pooling, dropout (0.5), and dense classification layers. Trained from scratch for 20 epochs.

3. **Transfer Learning (Head-Only)**: MobileNetV2 pre-trained on ImageNet as frozen feature extractor, custom classification head with global average pooling, dropout (0.3), and dense output layer. Trained for 10 epochs.
4. **Transfer Learning (Fine-Tuned)**: Same architecture with last 20 MobileNetV2 layers unfrozen for end-to-end fine-tuning. Trained for 5 additional epochs with reduced learning rate.

---

4. Results

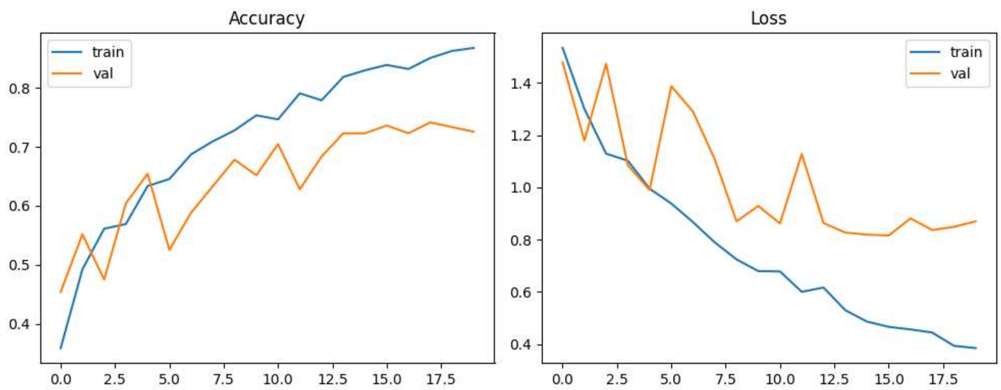
4.1 Model Performance Comparison

Model	Test Accuracy	Improvement
Logistic Regression	35.00%	-
Baseline CNN	69.21%	+34.21%
Transfer (Head-Only)	**83.42%**	+48.42%
Transfer (Fine-Tuned)	82.89%	+47.89%

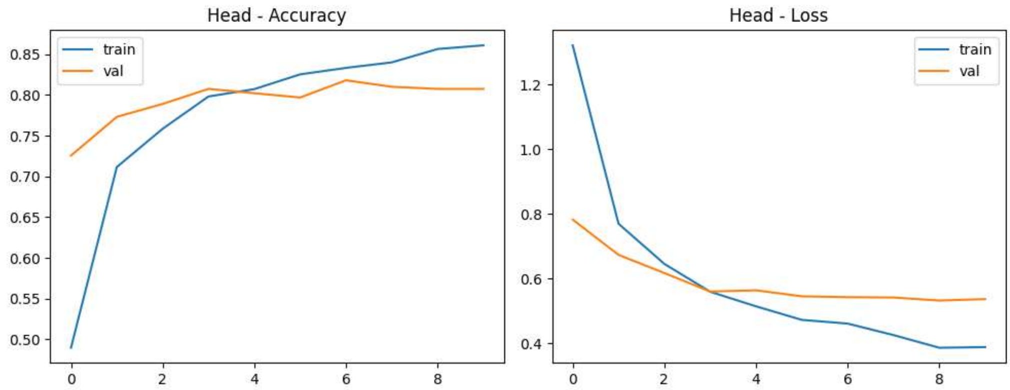
Table 1: Model comparison on test set (380 images)

Logistic Regression: Simple linear baseline operating on flattened 64×48 grayscale images achieved 35.00% test accuracy, significantly above random guessing (16.7%) but demonstrating the limitations of linear models on image data.

Baseline CNN: Custom architecture trained from scratch achieved 69.21% test accuracy after 20 epochs. Training accuracy reached 86.48% while validation plateaued at 74.14%, indicating significant overfitting despite dropout regularization. The 17.27 percentage point gap between training and test accuracy demonstrates the difficulty of learning effective features from scratch with limited data.



Transfer Learning (Head-Only): Using frozen MobileNetV2 features with a trainable classification head achieved 83.42% test accuracy — a remarkable 15.79 percentage point improvement over the baseline. This dramatic improvement demonstrates the power of pre-trained ImageNet features for waste classification with limited training data (1,768 images).



Transfer Learning (Fine-Tuned): Fine-tuning the last 20 MobileNetV2 layers for 5 epochs improved training accuracy to 89.48% but decreased test accuracy to 82.89%. This counterintuitive result suggests overfitting to the small training set, making the head-only model the optimal choice for deployment.

4.2 Detailed Performance Analysis

The best model (head-only transfer learning) achieved the following per-class results on the test set:

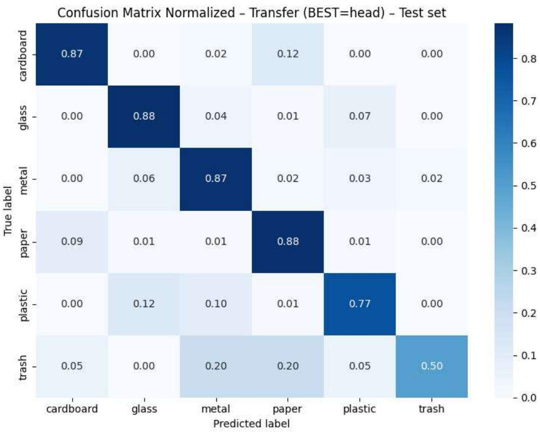
Class	**Precision**	**Recall**	**F1-Score**	**Support**
Cardboard	0.8525	0.8667	0.8595	60
Glass	0.8272	0.8816	0.8535	76
Metal	0.7714	0.8710	0.8182	62
Paper	0.8478	0.8764	0.8619	89
Plastic	0.8615	0.7671	0.8116	73
Trash	0.9091	0.5000	0.6452	20
Weighted Avg	0.8378	0.8342	0.8316	380

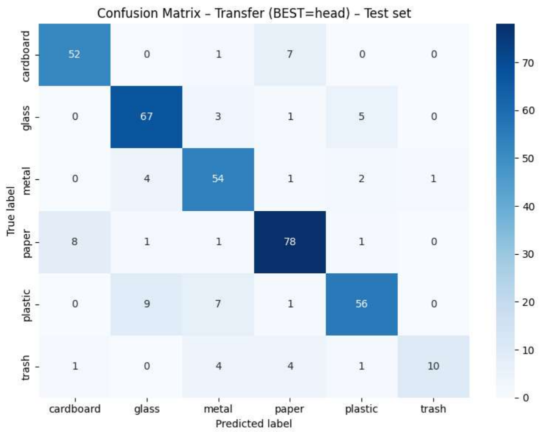
Table 2: Per-class performance metrics (Head-Only model)

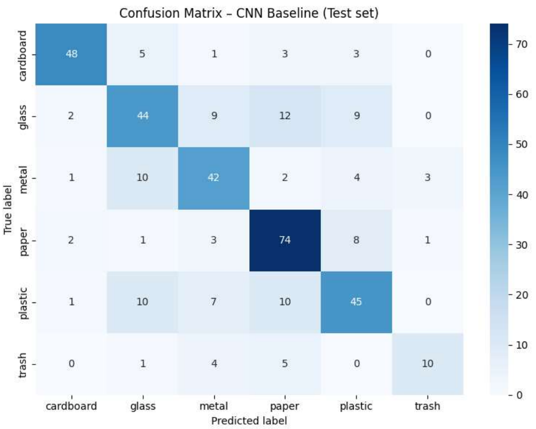
Key observations:

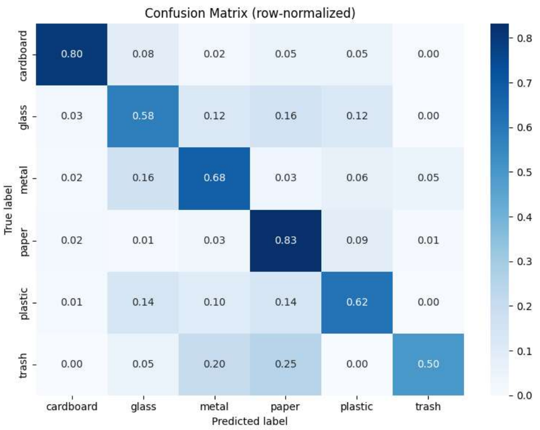
- **Best performance**: Paper (F1=0.8619) and Cardboard (F1=0.8595) achieved highest scores due to distinctive visual features, followed closely by Glass (F1=0.8535)

- **Challenging classes**: Trash (F1=0.6452) and Plastic (F1=0.8116) showed lower performance
- **Class imbalance impact**: Trash category (only 20 test samples) achieved highest precision (0.9091) but lowest recall (0.5000), indicating the model is very conservative in predicting this class
- **Trash category ambiguity**: The "trash" category is inherently ambiguous as it contains items that could be misclassified as other categories (e.g., damaged plastic items, contaminated paper), making it particularly challenging for the model to learn distinctive features
- **Confusion patterns**: Metal achieved high recall (0.8710) but lower precision (0.7714), suggesting some confusion with similar-looking metallic surfaces in other categories
- **Overall strong performance**: Glass improved significantly with 0.8816 recall, showing the model learned to distinguish transparent/reflective materials effectively









4.3 Real-World Generalization

We evaluated the head-only model on 62 custom photographs taken by the team under varied real-world conditions. While detailed metrics require further analysis, qualitative observations revealed:

- **Lighting sensitivity**: Performance degraded with non-standard lighting conditions
- **Background effects**: Cluttered backgrounds reduced classification confidence compared to clean training images. Notably, the brown wooden background used in many photographs likely caused confusion with the cardboard category, leading to misclassifications
- **Category-specific robustness**: Paper and cardboard maintained better performance; trash and plastic showed increased confusion
- **Domain shift**: The gap between controlled training data and real-world deployment scenarios indicates need for more diverse training examples with varied backgrounds

4.4 Discussion

Strengths: Transfer learning with MobileNetV2 achieved 83.42% accuracy using only 1,768 training images, demonstrating highly data-efficient learning with a 48.42 percentage point improvement over the logistic regression baseline (35.00%). The progression from linear classifier to CNN (+34.21%) to transfer learning (+14.21% additional) clearly demonstrates the value of spatial feature learning and pre-trained representations. The head-only training approach proved more effective than fine-tuning for this dataset size.

Limitations: Class imbalance particularly affects the trash category (96 training samples vs 416 for paper). The trash category is inherently ambiguous, containing items that overlap with other categories (damaged plastic, contaminated paper), making classification more difficult. Fine-tuning decreased test performance, suggesting the model overfits with limited data. Real-world generalization remains challenging due to domain shift, particularly background color effects (brown backgrounds confused with cardboard).

Recommendations: Expand training data especially for underrepresented classes, implement advanced augmentation techniques, consider ensemble methods, evaluate alternative architectures (EfficientNet, ResNet50), and deploy with human-in-the-loop verification for low-confidence predictions. The 83% accuracy supports assisted sorting applications but requires human oversight for production deployment.

---

5. Conclusion

This project successfully developed and compared machine learning models for automated garbage classification from images. The transfer learning approach using MobileNetV2 achieved 83.42% test accuracy, dramatically outperforming both the logistic regression baseline (35.00%) and custom CNN (69.21%). The progressive modeling approach demonstrated clear improvements: linear classifier → CNN (+34.21%) → transfer learning (+14.21% additional). Head-only training proved superior to fine-tuning for this dataset size (1,768 training images), demonstrating that pre-trained ImageNet features provide excellent representations for waste classification with limited data.

The system classifies six waste categories (cardboard, glass, metal, paper, plastic, trash) with per-class F1-scores ranging from 0.6452 to 0.8619. Class imbalance significantly impacts minority classes, particularly the ambiguous trash category with only 96 training samples that contains items overlapping with other categories. Real-world evaluation on 62 custom photographs revealed sensitivity to lighting and background variations, with brown backgrounds particularly causing confusion with cardboard classification, indicating need for more diverse training data with neutral backgrounds.

Future work should focus on dataset expansion for underrepresented classes, advanced augmentation strategies, ensemble methods, and domain adaptation techniques. With 83% accuracy, the system is suitable for assisted sorting applications where human operators verify low-confidence predictions. This work demonstrates that practical waste classification systems can be developed with modest datasets through effective transfer learning, contributing to automated waste management and sustainability initiatives.