Garbage Classification using EfficientNet-B3 and Deep Learning

Ankush Chakraborty

Department of Computer Science and Business Systems
Institute of Engineering and Management, Kolkata

Abstract. This paper presents a deep learning-based automatic garbage classification system using the EfficientNet-B3 architecture. The dataset, sourced from the Kaggle Garbage Classification V2 repository, contains 19,762 images across ten waste categories: metal, glass, biological, paper, battery, trash, cardboard, shoes, clothes and plastic. Due to class imbalance and high intra-class visual variation, the model was trained using transfer learning, strong data augmentation, weighted cross-entropy loss and selective fine-tuning of deeper EfficientNet layers. The proposed approach achieved a validation accuracy of 96.89%, with macro precision, recall and F1-scores of 95.85%, 96.51% and 96.16% respectively. These results demonstrate the suitability of EfficientNet-B3 for real-world waste recognition tasks and smart waste-management applications.

Keywords: EfficientNet-B3 · Garbage Classification · Deep Learning · Transfer Learning · Image Recognition

1 Introduction

Rapid urbanization and increasing consumerism have led to a significant rise in solid waste generation. Improper waste segregation contributes to environmental pollution, inefficient recycling and long-term ecological damage. Manual sorting of garbage is labor-intensive and prone to human error. Automated image-based garbage classification is therefore an important step toward smart waste-management systems.

Deep learning, particularly convolutional neural networks (CNNs), has shown excellent performance in image recognition tasks. However, standard CNN architectures may be computationally expensive or fail to generalize well on relatively small and imbalanced datasets. In this work, we employ EfficientNet-B3, a modern CNN architecture that achieves a good balance between accuracy and efficiency, and adapt it for a ten-class garbage classification problem.

The main contributions of this work are:

- Application of EfficientNet-B3 with transfer learning for multiclass garbage classification.
- Handling of dataset imbalance using class-weighted cross-entropy loss.
- Use of strong data augmentation and fine-tuning of the last EfficientNet blocks to reduce overfitting.

– A detailed experimental evaluation achieving 96.89% validation accuracy and strong macro F1-score.

2 Related Work

Several studies have explored deep learning for waste classification. Early approaches used traditional CNNs such as VGG16, AlexNet and GoogLeNet for binary recyclable/non-recyclable classification. Later works adopted architectures like ResNet, Inception and MobileNet to classify multiple categories of trash. While these models yield promising results, they often require large datasets and high computational resources.

EfficientNet, introduced by Tan and Le, proposes compound scaling of depth, width and input resolution, resulting in architectures that are both accurate and parameter-efficient. EfficientNet-B3, in particular, provides strong performance on image classification benchmarks with fewer parameters compared to many classical CNNs. Motivated by this, adopt EfficientNet-B3 as the backbone model for the garbage classification task.

3 Dataset Description

The experiments in this paper use the *Garbage Classification V2* dataset available on Kaggle.[1] The dataset consists of 19,762 RGB images belonging to ten categories of garbage.

Table 1. Class-wise distribution of the Garbage Classification V2 dataset.

Class	Number of Images
Metal	1020
Glass	3061
Biological	997
Paper	1680
Battery	944
Trash	947
Cardboard	1825
Shoes	1977
Clothes	5327
Plastic	1984
Total	**19,762**

As shown in Table 1, the dataset is imbalanced, with the *clothes* class having significantly more images than classes such as *biological*, *battery* and *trash*. This imbalance can bias the model toward majority classes if not addressed properly.

[1] `https://www.kaggle.com/datasets/sumn2u/garbage-classification-v2`

4 Proposed Methodology

4.1 Data Preprocessing and Augmentation

All images are resized to 300×300 pixels, which matches the recommended resolution for EfficientNet-B3. The pixel values are normalized using the ImageNet mean and standard deviation. To improve generalization and robustness, the following augmentation operations are applied during training:

- Random horizontal flip
- Random rotation (up to 20 degrees)
- Color jitter (brightness, contrast and saturation)
- Random perspective distortion
- Random erasing

These augmentations help the network become invariant to variations in orientation, lighting and partial occlusion.

4.2 Model Architecture

EfficientNet-B3 with ImageNet pretrained weights is used as the base feature extractor. Initially, all layers of EfficientNet-B3 are frozen. The original classifier head is replaced by a custom fully connected head consisting of:

- Dropout (0.5)
- Linear layer (1536 → 1024), ReLU activation, Batch Normalization
- Dropout (0.5)
- Linear layer (1024 → 512), ReLU activation, Batch Normalization
- Dropout (0.4)
- Linear layer (512 → 10) with softmax at inference time

After initially training only the classifier head, the final two MBConv blocks of EfficientNet-B3 are unfrozen and jointly fine-tuned with the classifier. This strategy leverages both general ImageNet features and task-specific features for garbage classification.

4.3 Class-Weighted Loss

Let n_c denote the number of samples in class c. To compensate for class imbalance, class weights w_c are computed as

$$w_c = \frac{1}{n_c}$$

and normalized. These weights are used in the cross-entropy loss function so that errors on minority classes contribute more strongly to the loss.

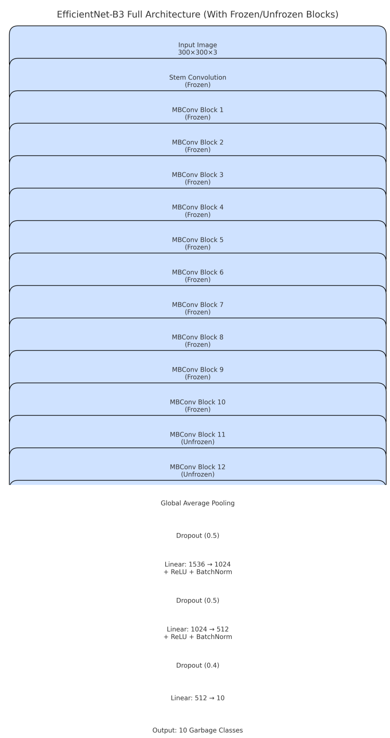

Fig. 1. High-level architecture of the proposed EfficientNet-B3 based garbage classifier, showing frozen and fine-tuned blocks.

4.4 Training Setup

The dataset is split into 70% training, 20% validation and 10% testing. The model is trained using the AdamW optimizer with an initial learning rate of 1×10^{-3} and weight decay of 1×10^{-4}. A Cosine Annealing learning-rate scheduler with $T_{\max} = 20$ is employed to gradually reduce the learning rate over epochs. The batch size is set to 32. Early stopping with a patience of 7 epochs is used to prevent overfitting.

5 Experimental Results

5.1 Validation Metrics

The model is evaluated on the validation split after each epoch. The best performance is observed at epoch 10 with the following metrics:

- Validation accuracy: 96.89%
- Macro precision: 95.85%
- Macro recall: 96.51%
- Macro F1-score: 96.16%

Table **??** summarizes the overall validation performance. The high macro F1-score indicates that the model performs well across all classes, not only the majority ones.

5.2 Discussion

The results demonstrate that EfficientNet-B3 is highly effective for garbage classification. Freezing the majority of the network and fine-tuning only the last two blocks strikes a balance between computational cost and model flexibility. The cosine annealing scheduler and strong data augmentation further contribute to stable training and reduced overfitting.

6 Conclusion and Future Work

This paper presented an EfficientNet-B3 based deep learning model for multiclass garbage classification. By combining transfer learning, class-weighted loss and selective fine-tuning, the model achieved a validation accuracy of 96.89% and a macro F1-score of 96.16% on the Garbage Classification V2 dataset. These results indicate that the approach is suitable for deployment in smart waste bins, automated sorting systems and other waste-management applications.

Future work will focus on expanding the dataset, improving performance on minority classes, incorporating object detection to localize garbage items within images and deploying the model on edge devices for real-time operation.

References

1. Sumn2u: Garbage Classification V2 Dataset. Kaggle (2020). https://www.kaggle.com/datasets/sumn2u/garbage-classification-v2
2. Tan, M., Le, Q.: EfficientNet: Rethinking model scaling for convolutional neural networks. In: Proceedings of the 36th International Conference on Machine Learning (ICML). pp. 6105–6114 (2019).
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016). .