Comparative Study on Binary Waste Classification Based on Deep Convolutional Neural Networks and Data Augmentation

Shuyuan Xing[a]
Information of Technology, The University of New South Wales, Sydney, 1466, Australia

Keywords: Waste Classification, Deep Convolutional Neural Networks, Data Augmentation.

Abstract: With the acceleration of urbanization and the increasing demand for environmental protection, waste classification has emerged as a crucial component of waste management. This paper proposes three baseline methods for binary waste classification based on deep convolutional neural networks and data augmentation techniques. The first baseline employs a pre-trained ResNet50 model combined with an SE attention module to enhance feature representation; the second baseline utilizes a lightweight EfficientNet-B0 model with conventional data augmentation strategies; and the third baseline also adopts EfficientNet-B0 but integrates more aggressive augmentation methods, such as random cropping, color jittering, Gaussian blur, and random erasing, to improve model generalization. Results from experiments on a Kaggle trash categorization dataset show that the EfficientNet-B0-based method with aggressive data augmentation significantly increases accuracy and robustness. This paper serves as a helpful reference for further research in this area since it not only presents an efficient deep learning solution for waste classification, but it also provides insightful information about how data augmentation techniques affect model performance.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

Globally, waste classification has become a crucial concern for resource recycling and environmental preservation. The amount of waste produced has increased due to fast urbanization and industrialization, making manual sorting techniques ineffective and prone to mistakes. As a result, they are unable to satisfy the growing waste management needs of contemporary cities. In addition to maximizing resource recovery and lowering pollution levels, efficient waste classification is crucial for promoting the circular economy by facilitating the recycling of valuable materials and lowering dependency on raw materials. Furthermore, proper waste classification can promote sustainable consumption and disposal behaviors by increasing recycling efficiency, reducing production costs, and increasing public environmental awareness.

Several technological approaches have been put forth in recent years to increase the effectiveness of trash classification. Convolutional Neural Networks (CNNs), in particular, have shown themselves to be effective tools for effectively classifying garbage photos and extracting rich feature information from them, greatly increasing identification accuracy (Ramsurrun et al., 2021). Using cutting-edge computer vision techniques, recent research has also investigated two-stage recognition-retrieval strategies to increase trash categorization accuracy (Zhang, S. et al., 2021). To further improve classification performance while reducing the requirement for sizable, labeled datasets, researchers have also used CNN optimization techniques and transfer learning (Zhang, Q. et al., 2021). By combining CNNs with deep feature refinement techniques, several studies have investigated hybrid systems and shown increased classification accuracy for intricate waste categories (Lu and Chen, 2022). Furthermore, smart city trash management systems now incorporate multi-agent simulations and Internet of Things (IoT) technology, enabling automated and more effective sorting procedures (Hussain et al., 2024). To improve real-time waste tracking and maximize resource allocation, smart waste management solutions also make use of IoT-based models (Mookkaiah et al., 2022).

[a] https://orcid.org/0009-0001-7677-6011

Xing, S.
Comparative Study on Binary Waste Classification Based on Deep Convolutional Neural Networks and Data Augmentation.
DOI: 10.5220/0013698300004670
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE 2025)*, pages 404-408
ISBN: 978-989-758-765-8
Proceedings Copyright © 2025 by SCITEPRESS – Science and Technology Publications, Lda.

This study explores the ways in which technological advancements can improve trash classification systems' precision and effectiveness. Our study intends to improve current approaches and present innovative ideas to address the problems with the waste sorting systems that are already in use by utilizing CNN-based image identification. This will ultimately help to create a more sustainable waste management strategy.

2 METHOD

According to recent research, aggressive data augmentation can greatly increase model robustness in real-world waste classification settings, especially when combined with domain adaptation strategies (Ruiz et al., 2019). Similar results have been found in earlier research, which showed that deep architectures like ResNet frequently overfit small datasets, requiring the use of extra regularization techzniques such data augmentation or dropout (Ogundana et al., 2024). Additionally, it has been demonstrated that using pre-trained deep learning models, like ResNet and EfficientNet, improves classification performance while lowering computing costs (Mao et al., 2021).

2.1 Data Preprocessing and Dataset Splitting

Three baseline methods were created in this study to deal with the binary waste classification issue. Despite having comparable basic functions, the baselines differed in the model architecture they used and the degree of data augmentation they used in preprocessing. The general framework, including data pretreatment, dataset partitioning, model creation, and training and evaluation techniques, is described in detail in the parts that follow.

Image data were first preprocessed using a number of common transformations for each of the three baselines. All photos in Baselines 1 and 2 were resized to 224 x 224 pixels as part of the preprocessing pipeline. This was followed by random rotation (up to 15°), random horizontal flipping, and color jittering (changing contrast and brightness) to add unpredictability. After that, images were transformed into tensors and normalized using the ImageNet dataset's mean and standard deviation. In order to further improve the classifier's resilience, a more aggressive augmentation method was used for Baseline 3. In addition to the usual random horizontal flip and rotation, this method included a randomly scaled crop to 224×224 with customizable scale and aspect ratio. Additionally, Baseline 3 used random erasing to mimic occlusions, Gaussian blur with a changeable sigma, and more noticeable color jittering (including saturation and hue adjustments).

2.2 Model Architectures

Throughout the baselines, two main deep learning architectures were used. A ResNet50 model that has already been trained on ImageNet is used in Baseline 1. In this configuration, the original final fully connected layer of ResNet50 is replaced by a new linear layer with a single output neuron to accommodate binary classification. An optional Squeeze-and-Excitation (SE) module is integrated into the network to recalibrate channel-wise feature responses and enhance the discriminative capacity of the extracted features.

In contrast, Baseline 2 and Baseline 3 adopt the EfficientNet-B0 architecture, which is known for its efficiency and scalability. Similar to the ResNet50 modification, the final classifier layer of EfficientNet-B0 is replaced with a linear layer producing a single output. Although both baselines use EfficientNet-B0 as the backbone, Baseline 2 applies the standard data augmentation strategy (similar to Baseline 1), while Baseline 3 incorporates the enhanced, aggressive augmentation pipeline described above. In both EfficientNet-based models, an optional SE module can be included, though the default configuration did not enable attention in our experiments.

2.3 Training and Evaluation Procedures

The Adam optimizer was used for training, with an initial learning rate of 1×10^{-4} for each baseline. The loss function was defined as Binary Cross Entropy with Logits (BCEWithLogitsLoss), which is appropriate for binary classification tasks since it naturally blends a Sigmoid activation with cross-entropy loss. The model's performance on the training set was tracked by computing the average loss and accuracy over the 50 epochs that the training process was iterated over. After each epoch, the model was evaluated on the validation set to assess generalization performance.

A progress bar that showed real-time loss numbers was used to monitor training progress. Furthermore, a logging mechanism was put in place to capture the model architecture specifics as well as the accuracy and loss metrics at each epoch to a log file. To give an objective assessment of the model's performance, it was further tested on a different test set after training.

The training and validation metrics were subsequently visualized using line plots to illustrate convergence behavior and model stability.

3 RESULTS

3.1 Baseline 1: ResNet50 + SE

Rapid convergence on the training set is indicated by the Baseline 1 (ResNet50 with the SE attention module) training and validation curves, which are displayed in figures 1 and 2. An outstanding fit to the training data is shown by the training accuracy reaching and remaining above 99% and the training loss rapidly decreasing to almost nil.

In contrast, the validation loss fluctuates between approximately 0.15 and 0.25, with the validation accuracy stabilizing around 93%–95%. These observations suggest that while the model is highly effective on the training set, there is some degree of overfitting. The final test results for Baseline 1 are Test Loss = 0.2858 and Test Accuracy = 93.63%, indicating that the ResNet50 + SE configuration is capable of robust feature extraction and generalizes well to unseen data despite the slight overfitting.

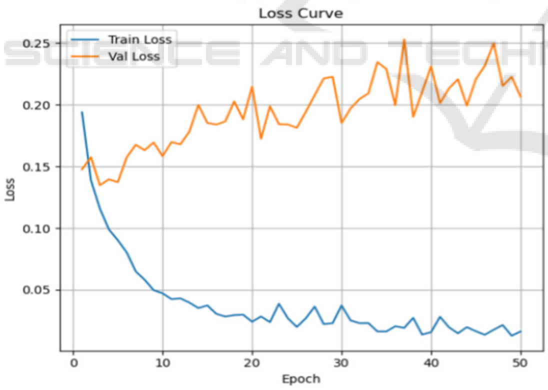

Figure 1: ResNet50 + SE Loss Curve. (Picture credit: Orignial)

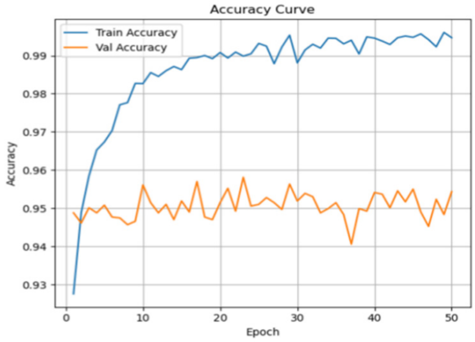

Figure 2: ResNet50 + SE Accuracy Curve. (Picture credit: Orignial)

3.2 Baseline 2: EfficientNet-B0 + Standard Data Augmentation

Baseline 2 employs the lightweight EfficientNet-B0 model along with a standard data augmentation strategy similar to that of Baseline 1. The training curves for Baseline 2 show rapid convergence, with the training loss quickly approaching zero and the training accuracy nearing 100%. However, the validation loss remains consistently higher, oscillating between 0.13 and 0.22, while the validation accuracy stays in the range of 94%–95% shown as figure 3 and 4. The final test results for Baseline 2 are Test Loss = 0.4420 and Test Accuracy = 93.16%. These metrics suggest that, although EfficientNet-B0 provides a faster and more lightweight alternative, its ability to generalize to the test set is slightly compromised when only conventional data augmentation is applied.

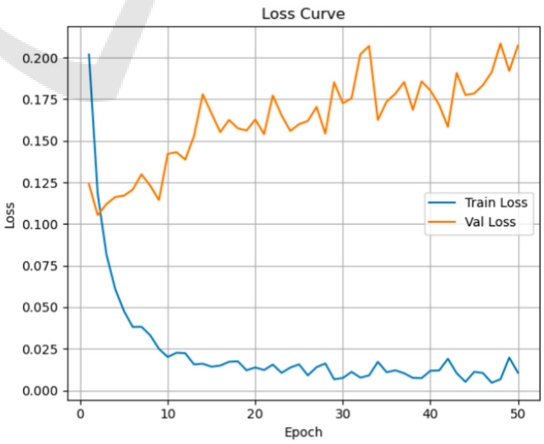

Figure 3: EfficientNet-B0 + Standard Data Augmentation Loss Curve. (Picture credit: Orignial)

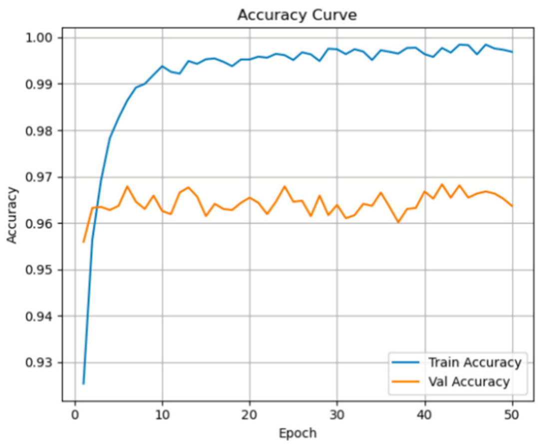

Figure 4: EfficientNet-B0 + Standard Data Augmentation Accuracy Curve. (Picture credit: Orignial)

3.3 Baseline 3: EfficientNet-B0 + Aggressive Data Augmentation

In Baseline 3, the same EfficientNet-B0 backbone is used; however, a more aggressive data augmentation pipeline is employed. This pipeline includes random resized cropping, stronger color jitter (with adjustments in saturation and hue), Gaussian blur, and random erasing. The training curves again show rapid convergence with the training loss approaching zero and near-perfect training accuracy. The validation loss for Baseline 3 remains within a similar range as in the previous baselines (approximately 0.15–0.25), and the validation accuracy is observed to be around 96% shown as figure 5 and 6. The final test results are Test Loss = 0.2964 and Test Accuracy = 93.51%, which are very close to those of Baseline 1 and slightly better than Baseline 2. These results demonstrate that the incorporation of aggressive data augmentation significantly enhances the generalization ability of the lightweight EfficientNet-B0 model, bringing its performance close to that of the deeper ResNet50 + SE model.

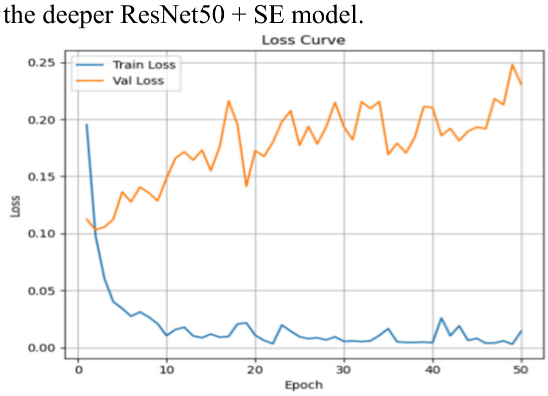

Figure 5: EfficientNet-B0 + Aggressive Data Augmentation Loss Curve. (Picture credit: Orignial)

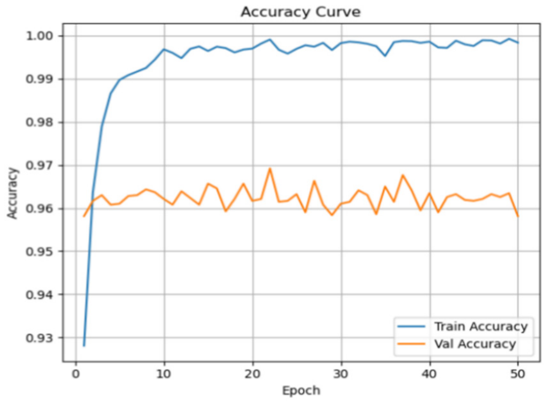

Figure 6: EfficientNet-B0 + Aggressive Data Augmentation Accuracy Curve. (Picture credit: Orignial)

3.4 Result Analysis

The paper designed and implemented three baseline methods for binary waste classification, employing a pre-trained ResNet50 with an SE attention module, an EfficientNet-B0 model with standard data augmentation, and an EfficientNet-B0 model enhanced with aggressive data augmentation. Experimental results demonstrated that all models converged rapidly during training, achieving nearly 100% training accuracy; on the test and validation sets, overfitting was noted, though. With a test accuracy of 93.63% and a test loss of 0.2858, the ResNet50 + SE model in particular demonstrated that deep networks and attention mechanisms may work together to efficiently extract picture data and perform well in waste classification tasks. In contrast, the EfficientNet-B0 model with standard data augmentation reached only 93.16% accuracy and a higher test loss of 0.4420, reflecting slightly weaker generalization performance. Significantly, the EfficientNet-B0 model's ability to adapt to a variety of image data was enhanced by implementing aggressive data augmentation strategies. It achieved a test accuracy of 93.51% and a significantly lower test loss of 0.2964, almost matching the ResNet50 + SE model's performance while providing advantages in terms of model size and computational efficiency, as indicated in table 1.

Table 1: Test result of three baseline model.

	Loss	Accuracy
ResNet50 + SE	0.2858	93.63%
EfficientNet-B0 + Standard Data Augmentation	0.4420	93.16%
EfficientNet-B0 + Aggressive Data Augmentation	0.2964	93.51%

4 CONCLUSIONS

Overall, our study demonstrates that while deep architectures like ResNet50 + SE are capable of extracting rich image features and achieving high classification accuracy, lightweight models such as EfficientNet-B0 can attain comparable performance when enhanced with advanced data augmentation techniques, making them particularly suitable for resource-constrained applications. Data augmentation is shown to play a critical role in improving model generalization by effectively mitigating overfitting and ensuring robust performance on unseen data.

In order to further minimize overfitting, future research will investigate more complex regularization techniques and dynamic learning rate scheduling. To reduce overfitting in trash categorization models, several optimization strategies have been put forth, including adaptive weight decay and dynamic learning rate scheduling. Furthermore, it is anticipated that adding attention mechanisms or other sophisticated feature reconstruction methods to lightweight models may improve classification performance even further without requiring a large amount of processing cost. To improve the robustness and generalization of waste classification systems in complicated circumstances, efforts will also be made to expand and refine the dataset, explore semi-supervised or unsupervised learning methodologies, and use cross-domain data fusion techniques. In the end, these study developments will offer technical assistance for the implementation of automated waste sorting systems that is more effective, economical, and long-lasting.

REFERENCES

Hussain, I., Elomri, A., Kerbache, L., and El Omri, A. 2024. Smart city solutions: Comparative analysis of waste management models in IoT-enabled environments using multiagent simulation. Sustainable Cities and Society, 103, 105247. doi: 10.1016/j.scs.2024.105247.

Jin, S., Yang, Z., Królczyk, G., Liu, X., Gardoni, P., and Li, Z. 2023. Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling. Waste Management, 162, 123–130. doi: 10.1016/j.wasman.2023.02.014.

Lu, W., and Chen, J. 2022. Computer vision for solid waste sorting: A critical review of academic research. Waste Management, 142, 29–43. doi: 10.1016/j.wasman.2022.02.009.

Mao, W.-L., Chen, W.-C., Wang, C.-T., and Lin, Y.-H. 2021. Recycling waste classification using optimized convolutional neural network." Resources, Conservation and Recycling, 164, 105132. doi: 10.1016/j.resconrec.2020.105132.

Mookkaiah, S. S., Thangavelu, G., Hebbar, R., et al. 2022. "Design and development of smart Internet of Things–based solid waste management system using computer vision." Environmental Science and Pollution Research, 29, 64871–64885. doi: 10.1007/s11356-022-20428-2.Goals (SEB4SDG), Omu-Aran, Nigeria, 1–11. doi: 10.1109/SEB4SDG60871.2024.10629933.

Ogundana, A. K., Afolabi, O. O., Ilevbare, M., and Falae, P. O. (2024). "Green Hydrogen Generation from Plastic Waste: A Review." In 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 1–11. doi: 10.1109/SEB4SDG60871.2024.10629933.

Ramsurrun, N., Suddul, G., Armoogum, S. and Foogooa, R. 2021. Recyclable Waste Classification Using Computer Vision And Deep Learning. in 2021 Zooming Innovation in Consumer Technologies Conference (ZINC), Novi Sad, Serbia, pp. 11–15, doi: 10.1109/ZINC52049.2021.9499291.

Ruiz, V., Sánchez, Á., Vélez, J. F., and Raducanu, B. 2019. "Automatic Image-Based Waste Classification." In From Bioinspired Systems and Biomedical Applications to Machine Learning. IWINAC 2019, Lecture Notes in Computer Science, 11487, edited by Ferrández Vicente, J., Álvarez-Sánchez, J., de la Paz López, F., Toledo Moreo, J., and Adeli, H. Springer, Cham. doi: 10.1007/978-3-030-19651-6_41.

Zhang, Q., Yang, Q., Zhang, X., Bao, Q., Su, J., and Liu, X. 2021. Waste image classification based on transfer learning and convolutional neural network. Waste Management, 135, 150–157. doi: 10.1016/j.wasman.2021.08.038.

Zhang, S., Chen, Y., Yang, Z., and Gong, H. 2021. Computer Vision Based Two-stage Waste Recognition-Retrieval Algorithm for Waste Classification. Resources, Conservation and Recycling, 169, 105543. doi: 10.1016/j.resconrec.2021.105543.